

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 8, 2003, 05:14:09 ; Search time 37.454 Seconds
(without alignments)
1359.222 Million cell updates/sec

Title: US-09-813-775c-18

Perfect score: 843
Sequence: 1 APPRLICDSRVLEERYLEAK.....NFLRGLKLYTGACRTGDR 166

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+pn.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09813777/rnual.07012003.153121.23808/app-query.fasta_1.718
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOORCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09813775.@CGN_1_18.@rnual.07012003.153121.23808 -MCP=6 -ICPU=3
-NO_XLPHY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

- Issued_Patents_NA:*
- 1: /cgn2_6/plodata/1/lna/5A.COMB.seq:*
 - 2: /cgn2_6/plodata/1/lna/5B.COMB.seq:*
 - 3: /cgn2_6/plodata/1/lna/6A.COMB.seq:*
 - 4: /cgn2_6/plodata/1/lna/6B.COMB.seq:*
 - 5: /cgn2_6/plodata/1/lna/PCTUS.COMB.seq:*
 - 6: /cgn2_6/plodata/1/lna/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	825	97.9	508	1	US-08-318-193-69 Sequence 69, Appl
2	825	97.9	725	1	PCT-US94-04361-21 Sequence 21, Appl
3	825	97.9	788	2	US-08-883-795A-35 Sequence 35, Appl
4	825	97.9	823	2	US-08-883-795A-35 Sequence 1, Appl
5	825	97.9	823	4	US-09-205-337-1 Sequence 12, Appl
6	825	97.9	1011	2	US-08-750-128-12 Sequence 39, Appl
7	822	97.5	1239	4	US-09-366-009-39 Sequence 7, Appl
8	802	95.1	945	2	US-08-750-128-4 Sequence 4, Appl
9	802	95.1	969	2	US-08-750-128-4 Sequence 3, Appl
10	743.5	88.2	5107	4	US-08-910-647-3 Sequence 3, Appl
11	743.5	88.2	5107	4	US-09-620-925-3 Sequence 22, Appl
12	741.5	88.0	681	5	PCT-US94-04361-22

13	695	82.4	681	5	PCT-US94-04361-27	Sequence 27, Appl
14	692	82.1	678	5	PCT-US94-04361-24	Sequence 24, Appl
15	676	80.2	679	5	PCT-US94-04361-23	Sequence 23, Appl
16	665.5	78.9	687	5	PCT-US94-04361-25	Sequence 25, Appl
17	665	78.9	688	5	PCT-US94-04361-26	Sequence 26, Appl
18	590.5	70.0	3601	3	US-09-017-631-23	Sequence 23, Appl
19	590.5	70.0	3602	2	US-08-883-795A-33	Sequence 33, Appl
20	590.5	70.0	3602	2	US-09-018-138-1	Sequence 1, Appl
21	106	12.6	65	2	US-08-883-795A-6	Sequence 24, Appl
22	105	12.5	58	2	US-08-883-795A-22	Sequence 6, Appl
23	105	12.5	58	2	US-08-883-795A-5	Sequence 22, Appl
24	97	11.5	56	2	US-08-883-795A-8	Sequence 5, Appl
25	97	11.5	61	2	US-08-883-795A-5	Sequence 8, Appl
26	96	11.4	61	2	US-08-883-795A-27	Sequence 27, Appl
27	96	11.4	996	4	US-08-875-533-59	Sequence 59, Appl
28	95	11.3	996	4	US-08-875-533-60	Sequence 60, Appl
29	93	11.0	459	4	US-08-875-533-58	Sequence 58, Appl
30	93	11.0	465	3	US-08-469-318-180	Sequence 180, Appl
31	93	11.0	465	3	US-08-468-609A-180	Sequence 180, Appl
32	93	11.0	465	3	US-08-471-045-54	Sequence 54, Appl
33	93	11.0	465	3	US-08-469-712A-54	Sequence 54, Appl
34	93	11.0	465	4	US-08-446-871-54	Sequence 54, Appl
35	93	11.0	465	4	US-08-446-872A-180	Sequence 180, Appl
36	93	11.0	465	4	US-08-468-910-54	Sequence 54, Appl
37	93	11.0	465	4	US-08-761-907-54	Sequence 54, Appl
38	93	11.0	465	4	US-08-762-227A-180	Sequence 180, Appl
39	93	11.0	465	5	PCT-US95-01185-180	Sequence 180, Appl
40	93	11.0	498	1	US-08-413-803-28	Sequence 28, Appl
41	93	11.0	605	1	US-08-388-779A-3	Sequence 3, Appl
42	93	11.0	605	1	US-08-591-070A-3	Sequence 3, Appl
43	93	11.0	605	2	US-08-927-855-3	Sequence 3, Appl
44	93	11.0	852	3	US-08-469-318-185	Sequence 185, Appl
45	93	11.0	852	3	US-08-468-609A-185	Sequence 185, Appl

ALIGNMENTS

RESULT 1
US-08-318-193-69
Sequence 69, Application US/08318193
Patent No. 5641663
GENERAL INFORMATION:
APPLICANT: GARVIN, Robert T.
TITLE OF INVENTION: MALEK, Lawrence T.
TITLE OF INVENTION: AN EXPRESSION SYSTEM FOR THE SECRETION
OF BIOACTIVE HUMAN GRANULOCYTE MACROPHAGE COLONY
STIMULATING FACTOR (GM-CSF) AND OTHER HETEROLOGOUS
PROTEINS FROM STREPTOCOCCI
NUMBER OF INVENTION: 91
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 diagonal Road, Suite 500
CITY: Alexandria
STATE: Virginia
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,193
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,314
FILING DATE:
APPLICATION NUMBER: US 07/224,568
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 18740/116 CACO

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703)836-9300
 TELEFAX: (703)683-4109
 TELEX: 899149
 INFORMATION FOR SEQ ID NO: 69:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 508 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: Other nucleic acid;
 DESCRIPTION: Synthetic DNA oligonucleotide
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 5..502
 US-08-318-193-69

Alignment Scores:
 Pred. No.: 5,06e-101 Length: 508
 Score: 825.00 Matches: 162
 Percent Similarity: 97.59% Conservative: 0
 Best Local Similarity: 97.59% Mismatches: 4
 Query Match: 97.86% Indels: 0
 DB: 1 Gaps: 0

US-09-813-775C-18 (1-166) x US-08-318-193-69 (1-508)

QY 1 AAlaProPtoArlgLeuIlleCysAspSerArqValLeuGluArqTyrLeuLeuGluAlaLys 20
 DB 5 GCCCCCCCGCGCTCATCTGCGACAGCCGCTCTCGAGCGGTACTGTCGACAGCCAAAG 64
 QY 21 GluAlaGluAsnIleThrThrlThrlGlySalGluHisCysSerLeuAsnGluAsnIleThr 40
 DB 65 GAGGGGGAATATACGACGGGGTGGCCGACACTGCTCCCTCACGAGAACATATCACCC 124
 QY 41 ValProAspThrLysValAsnPhetYrAlaTrpLysArgAsn***Ser***GlnGlnAla 60
 DB 125 GTCGCCGACACCAAGTCACTTACGCTGGAAGCCGATGAGGTGGCCACGACAGCGG 184
 QY 61 ValGluValITrPGlnGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
 DB 185 GTGAGAGTCTGGCAGGGGCTCGGCTCTCTCGAGGGGCTCTCGGGGACAGGCCCTCC 244
 QY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSer 100
 DB 245 CTGGTGAACCTGCTCCACGCGGTGGAGCGGCTCCAGCTGCACGACAAAGCCGCTCTCC 304
 QY 101 GlyLeuArgSerLeuThrlThrlLeuArgAlaLeuGlyAlaGlnLysGluAlaLysSer 120
 DB 305 GGGCTCCGCTCTACACACGCTGCTGGCGGCTCGGTGGCCACAGAGGACCATCTCG 364
 QY 121 ProProAspAlaAlaSerAlaAlaProLeuArgTrlIleThrlAlaAspThrPheArgLys 140
 DB 365 CCCCCGAGCGCGCAGCGCGCGCGCTGCGGACATCAGCGGAGACCTTCGCGCAAG 424
 QY 141 LeuPheArgValTyrSerAsnPhleuArgGlyLysLeuLysLeuTyrThrlGlyAla 160
 DB 425 CTGTTCCGGGCTCTACTCAACTCTCTGGGGGAGAGTGAACCTCTACACCGCGGAGGCC 484
 QY 161 CysArgThrGlyAspArg 166
 DB 485 TGCCGCGACGGGAGACCGG 502

RESULT 2
 PCT-US94-04361-21
 Sequence 21, Application PC/TUS9404361
 GENERAL INFORMATION:
 APPLICANT: Brigham and Women's Hospital
 APPLICANT: 75 Francis Street
 APPLICANT: Boston, MA 02115
 APPLICANT: Bunn, H. Franklin
 APPLICANT: Wen, Danyi
 APPLICANT: Showers, Mark O.

TITLE OF INVENTION: Erythropoietin Mutelins With Enhanced
 TITLE OF INVENTION: Activily
 NUMBER OF SEQUENCES: 59
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox
 STREET: 1100 New York Avenue, Suite 600
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005-3934
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentln Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/04361
 FILING DATE: Herewith
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/049,802
 FILING DATE: 21-APR-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Cimbala, Michele A.
 REGISTRATION NUMBER: 33,851
 REFERENCE NUMBER: 0627.336PC01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2540
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 725 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: both
 PCT-US94-04361-21

Alignment Scores:
 Pred. No.: 8.77e-101 Length: 725
 Score: 825.00 Matches: 162
 Percent Similarity: 97.59% Conservative: 0
 Best Local Similarity: 97.59% Mismatches: 4
 Query Match: 97.86% Indels: 0
 DB: 5 Gaps: 0

US-09-813-775C-18 (1-166) x PCT-US94-04361-21 (1-725)

QY 1 AAlaProPtoArlgLeuIlleCysAspSerArqValLeuGluArqTyrLeuLeuGluAlaLys 20
 DB 92 GCCCCACACGCGCTCATCTGTGACAGCCGAGTCTCGGAGAGTACTCTTGGAGGCCAAG 151
 QY 21 GluAlaGluAsnIleThrThrlThrlGlySalGluHisCysSerLeuAsnGluAsnIleThr 40
 DB 152 GAGGCGGAGAAATATACGACGGGCTGTGTGAACTGACACTGTAAGAGACATATCTCT 211
 QY 41 ValProAspThrLysValAsnPhetYrAlaTrpLysArgAsn***Ser***GlnGlnAla 60
 DB 212 GTCCGACAGACCAAGTAATTTCTATGCTGGAAGAGATGAGTGGCGACAGAGGCC 271
 QY 61 ValGluValITrPGlnGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
 DB 272 GTAGAAAGTCTGGCAGGGGCTGGCTGCTGTGGAAAGTGTCTGCGGGGCGAGCCCTG 331
 QY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSer 100
 DB 332 TTGGTCAACTCTTCCACACGCTGGAGGCCCTGCGACGTGATGATTAAGCCGTCACT 391
 QY 101 GlyLeuArgSerLeuThrlThrlLeuArgAlaLeuGlyAlaGlnLysGluAlaLysSer 120
 DB 392 GGGCTTCCACACCTTCACACTCTGCTGGGCTGTGGAGCCCGAAGAGGACCATCTCC 451
 QY 121 ProProAspAlaAlaSerAlaAlaProLeuArgTrlIleThrlAlaAspThrPheArgLys 140

DB 452 CCTCCAGATGGGCTGCTGCTCCACTCCGAACAATCATCTGCTGACACTTCCGCAA 511

OY 141 leupheargValtyrSeranpheleuArglyLysleuLysleuYrThrIgluAla 160
|||||
DB 512 CFTCTCCGAGTCTACTCAATTTCTCTCCGGGAAAGCTGAAGCTGTACACAGGGAGGCC 571

OY 161 CysArgThrIgluAspArg 166
|||||
DB 572 TGCAGCAGACGAGGACAGA 589

RESULT 3

US-08-883-795A-35

Sequence 35, Application US/08883795A

Patent No. 5985607

GENERAL INFORMATION:

APPLICANT: Delcove, Genevieve

APPLICANT: Awang, Gregor

TITLE OF INVENTION: Recombinant DNA Molecules and Expression

TITLE OF INVENTION: Vectors for Tissue Plasmidogen Activator

NUMBER OF SEQUENCES: 39

CORRESPONDENCE ADDRESS:

ADDRESSEE: BERESKIN & PARR

STREET: 40 King Street West

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5H 3Y2

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/883,795A

FILING DATE: 27-JUN-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Gravelle, Micheline

REGISTRATION NUMBER: 40,261

REFERENCE/DOCKET NUMBER: 7841-062

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 364-7311

TELEFAX: (416) 361-1398

INFORMATION FOR SEQ ID NO: 35:

SEQUENCE CHARACTERISTICS:

LENGTH: 788 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

IMMEDIATE SOURCE:

CLONE: EPOlong

US-08-883-795A-35

Alignment Scores:

Pred. No.: 9,98e-101 Length: 788

Score: 825.00 Matches: 162

Percent Similarity: 97.59% Conservative: 0

Best Local Similarity: 97.59% Mismatches: 4

Query Match: 97.86% Indels: 0

DB: 2 Gaps: 0

US-09-813-775C-18 (1-166) x US-08-883-795A-35 (1-788)

OY 1 AlaProProArgLeuIleCysAspSerArgValLeuGluArgTyrIleuLeuGluAlaLys 20
|||||
DB 161 GCCCCACACGGCTCATCTGTGACAGCCGAGTCTGAGAGGTACTCTTGGAGGCCAAG 220

OY 21 GluAlaGluAsnIleThrThrIgluLysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
|||||
DB 221 GAGGCCGAGATATTCACGACGGGCTGTGTAACATTGCAAGCTTGTAATGAGATATTCAC 280

OY 41 ValProAspThrLysValAsnPhetYrAlaTrpLysArgAsn***Ser***GlnGlnAla 60
|||||

DB 281 GTCCACAGACACCAAGTAAATTCTATGCTCGAGAGAGATGAGGTCGGGACGAGGCC 340

OY 61 ValGluValTrpIleGlnLysLeuAlaLeuLeuSerGluAlaValLeuArgIgluAlaLeu 80
|||||

DB 341 GTAGAAAGTGTGGCAGGGCCCTGGCCCTGCTGTCCGAAAGCTGTCTCGGGGGCCAGGCCCTG 400

OY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSer 100
|||||

DB 401 TTGGTCAACTCTTCCAGCGCTGGAGGCCCTCAGCTGATGTGATTAAGCGTCAGT 460

OY 101 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyValGlnLysGluAlaIleSer 120
|||||

DB 461 GGCCTTCGACGCTCACCACATCTGCTGGGCTCTGGAGACCCAGAAAGCCATCTCC 520

OY 121 ProProAspAlaAlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLys 140
|||||

DB 521 CCTCCAGATGGGCTCAGCTGCTCCACTCCGAACAATCATCTGCTGACACTTCCGCAA 580

OY 141 leupheargValtyrSeranpheleuArglyLysleuLysleuYrThrIgluAla 160
|||||

DB 581 CTCTCCGAGTCTACTCAATTTCTCTCCGGGAAAGCTGAAGCTGTACACAGGGAGGCC 640

OY 161 CysArgThrIgluAspArg 166
|||||

DB 641 TGCAGCAGACGAGGACAGA 658

RESULT 4

US-08-785-750-1

Sequence 1, Application US/08785750

Patent No. 5846528

GENERAL INFORMATION:

APPLICANT: PODSAKOFF, GREGORY M.

APPLICANT: KURTZMAN, GARY J.

TITLE OF INVENTION: METHODS OF TREATING ANEMIA USING

TITLE OF INVENTION: RECOMBINANT ADEMO-ASSOCIATED VIRUS VIRIONS

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: ROBINS & ASSOCIATES

STREET: 90 MIDDLEFIELD ROAD, SUITE 200

CITY: MENLO PARK

STATE: CA

COUNTRY: USA

ZIP: 94025

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/785,750

FILING DATE: 16-JAN-1997

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/588,355

FILING DATE: 18-JAN-1996

ATTORNEY/AGENT INFORMATION:

NAME: MCCracken, THOMAS P.

REGISTRATION NUMBER: 38,548

REFERENCE/DOCKET NUMBER: 0800-0009.21

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 325-7812

TELEFAX: (415) 325-7823

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 823 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-785-750-1

Alignment Scores:

Pred. No.: 1.07e-100 Length: 823
Score: 825.00 Matches: 162
Percent Similarity: 97.59% Conservative: 0
Best Local Similarity: 97.59% Mismatches: 4
Query Match: 97.86% Indels: 0
Gaps: 2

US-09-813-775c-18 (1-166) x US-08-785-750-1 (1-823)

OY 1 AAlaProProAArgLeuIleCysAspSerArgValLeuGluArgTyrLeuLeuGluAlaLys 20
Db 133 GCCCAGACAGCCGCTCATGTGTACAGCCGAGCTCGAGAGAGTACCTCTTGAGGCCAAG 192
OY 21 GluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
Db 193 GAGGCCGGAATATCATCAGCAGGCGCTGTGTGAACTGCACTGCACTTAATGAGATTCAC 252
OY 41 ValProAspThrLysValAsnPhetYrAlaTrpLysArgAsn***Ser***GlnGlnAla 60
Db 253 GTCCAGACACCAAGTAAATTTCTATGCTGAGAGAGATGAGTGGCAGCAGGCC 312
OY 61 ValGluValITrpgInGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
Db 313 GTAGAAGTCTGGCAGGCGCTGCTGTGTGAGAGCTGCTGCGGGCCAGGCCCTG 372
OY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSer 100
Db 373 TTGGTCAACTCTTCCAGCCGCGGAGCCCTGCACTGCTGATGATTAAGCCGCTCACT 432
OY 101 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGlnAlaIleSer 120
Db 433 GGCCTTCCAGCCTCAGCAGCTCTGCTGGGCTGCGGAGCCAGAGAGCCATCTCC 492
OY 121 PropProAspAlaIleSerAlaIleProLeuArgThrIleThrAlaAspThrPheArgLys 140
Db 493 CTTCCAGATGGGCTCAGCTGCTCCTCAGCAATCATCAGTCTGAGCACTTTCGCAAA 552
OY 141 LeuPheArgValITyrSerAsnPhetLeuArgGlyLysLeuLysLeuTyrThrGlyGlnAla 160
Db 553 CTCTCCGAGTCTACTCCAAATTTCTCCGCGGGAAGCTGAAGCTGTACACAGGGAGGCC 612
OY 161 CysArgThrGlyAspArg 166
Db 613 TGCAGACAGGAGACAGA 630

RESULT 5

US-09-205-337-1
Sequence 1, Application US/09205337
Patent No. 632598

GENERAL INFORMATION:

APPLICANT: PODSAKOFF, GREGORY M.
KURTZMAN, GARY J.

TITLE OF INVENTION: METHODS OF TREATING ANEMIA USING

RECOMBINANT ADENO-ASSOCIATED VIRUS VIRIONS

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESS: ROBINSON & ASSOCIATES

STREET: 90 MIDDLEFIELD ROAD, SUITE 200

CITY: MENLO PARK

STATE: CA

COUNTRY: USA

ZIP: 94025

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/205,337

FILING DATE: 04-Dec-1998

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/785,750

FILING DATE: <unknown>

ATTORNEY/AGENT INFORMATION:

NAME: MCCracken, THOMAS P.

REGISTRATION NUMBER: 38,548

REFERENCE/DOCKET NUMBER: 0800-0009.21

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 325-7812

TELEFAX: (415) 325-7823

INFORMATION FOR SEQ. ID NO. 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 823 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Alignment Scores:

Pred. No.: 1.07e-100 Length: 823
Score: 825.00 Matches: 162
Percent Similarity: 97.59% Conservative: 0
Best Local Similarity: 97.59% Mismatches: 4
Query Match: 97.86% Indels: 0
Gaps: 2

US-09-813-775c-18 (1-166) x US-09-205-337-1 (1-823)

OY 1 AAlaProProAArgLeuIleCysAspSerArgValLeuGluArgTyrLeuLeuGluAlaLys 20
Db 133 GCCCAGACAGCCGCTCATGTGTACAGCCGAGCTCGAGAGAGTACCTCTTGAGGCCAAG 192
OY 21 GluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
Db 193 GAGGCCGGAATATCATCAGCAGGCGCTGTGTGAACTGCACTGCACTTAATGAGATTCAC 252
OY 41 ValProAspThrLysValAsnPhetYrAlaTrpLysArgAsn***Ser***GlnGlnAla 60
Db 253 GTCCAGACACCAAGTAAATTTCTATGCTGAGAGAGATGAGTGGCAGCAGGCC 312
OY 61 ValGluValITrpgInGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
Db 313 GTAGAAGTCTGGCAGGCGCTGCTGTGTGAGAGCTGCTGCGGGCCAGGCCCTG 372
OY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSer 100
Db 373 TTGGTCAACTCTTCCAGCCGCGGAGCCCTGCACTGCTGATGATTAAGCCGCTCACT 432
OY 101 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGlnAlaIleSer 120
Db 433 GGCCTTCCAGCCTCAGCAGCTCTGCTGGGCTGCGGAGCCAGAGAGCCATCTCC 492
OY 121 PropProAspAlaIleSerAlaIleProLeuArgThrIleThrAlaAspThrPheArgLys 140
Db 493 CTTCCAGATGGGCTCAGCTGCTCCTCAGCAATCATCAGTCTGAGCACTTTCGCAAA 552
OY 141 LeuPheArgValITyrSerAsnPhetLeuArgGlyLysLeuLysLeuTyrThrGlyGlnAla 160
Db 553 CTCTCCGAGTCTACTCCAAATTTCTCCGCGGGAAGCTGAAGCTGTACACAGGGAGGCC 612
OY 161 CysArgThrGlyAspArg 166
Db 613 TGCAGACAGGAGACAGA 630

RESULT 6

US-08-750-128-12
Sequence 12, Application US/08750128
Patent No. 5916773

GENERAL INFORMATION:

APPLICANT: MELE Antonio,

APPLICANT: DE SANTIS Rita,

Query Match: 97.51% Indels: 0
DB: 4 Gaps: 0

US-09-813-775C-18 (1-166) x US-09-366-009-39 (1-1239)

QY 1 AAlaProPArgLeuIleIcysAspSerArgValLeuGluArgTyrLeuLeuGluAlaLys 20
DB 697 GCCCACCACCGCTCTGTCAGCAGCCAGTCTGCGAGGAGTACCTCTTGAGGCGCAAG 756
QY 21 GluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
DB 757 GAGCGCCGGAATATCATCAGCAGCGGCTGTGCTGAACACTGCAGCTTGATGAGAAATATCACT 816
QY 41 ValProAspThrLysValAlaAsnPhenylalanylArgAsn***Ser***GlnGlnAla 60
DB 817 GTCCAGACACCAAGATTATTTCTATGCTTGAGAGATGAGAGCTGCGGCGACGAGGCC 876
QY 61 ValGluValATrPGLngIleuValLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
DB 877 GTAGAAAGTCTGCGAGGCGCTGCGCTGCTGTCGGAAGCTCTCGCGGCGCAGGCGCTG 936
QY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAlaAspLysAlaValSer 100
DB 937 TTGCTCACTCTTCCAGCGCTGCGGAGGCCCTGCGAGCTGCATGTGATTAACCGCTCACT 996
QY 101 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGluAlaIleSer 120
DB 997 GCCCTTCCAGGCTCAGCAGCTCTGCTGCGGCTGCGGAGCCAGAGGAGCAATCTCC 1056
QY 121 PropioAspAlaAlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLys 140
DB 1057 CCTCAGATGCGGCGCTCAGCTCTCCACTCCGAAATCATCAGTCTGACACTTCCGCAAA 1116
QY 141 LeuPheArgValTyrSerAsnPhenylalanylLysLeuLysLeuTyrThrGlyGluAla 160
DB 1117 CTTCTCCAGTCTACTCCAAATTTCTCCGCGGAGAAAGCTGAAGCTGATACAGAGGAGGCC 1176
QY 161 CysArgThrGlyAspArg 166
DB 1177 TGCAGGACAGGGGACAGA 1194

RESULT 8
US-08-750-128-7
Sequence 7, Application US/08750128
Patent No. 5916773
GENERAL INFORMATION:
APPLICANT: MELE Antonio,
APPLICANT: DE SANTIS Rita,
APPLICANT: CARLONI Cristina,
APPLICANT: COSCARIELLA Annamaria
TITLE OF INVENTION: Hybrid molecule of formula GM-CSF-L-EPO or
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: LOWE, PRICE, LeBLANC & BECKER
STREET: 99 Canal Center Plaza, Suite 300
CITY: Alexandria
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk, 3.50 inch.
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/750,128
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IT FI 94 A 000106
FILING DATE: 27-MAY-1994
ATTORNEY/AGENT INFORMATION:

NAME: Robert L. Price
REGISTRATION NUMBER: 22,685
REFERENCE/DOCKET NUMBER: 2879-007
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-684-1111
TELEFAX: 703-684-1124
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 945 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHEetical: NO
US-08-750-128-7

Alignment Scores:

Pred. No.:	1-55e-97	Length:	945
Score:	802.00	Matches:	158
Percent Similarity:	97.53%	Conservative:	0
Best Local Similarity:	97.53%	Mismatches:	4
Query Match:	95.14%	Indels:	0
DB:	2	Gaps:	0

US-09-813-775C-18 (1-166) x US-08-750-128-7 (1-945)

QY 5 LeuIleCysAspSerArgValLeuGluArgTyrLeuLeuGluAlaLysGluAlaGluAsn 24
DB 457 CTCATCTGTGACAGCGGAGTCTCTGAGAGAGTACTCTTGAGGCGCAAGAGCGCAGAAAT 516
QY 25 ILeuThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThrValProAspThr 44
DB 517 ATCAGAGCGGGCTGTGCTGGAACATGCACTGGAATGAAATATCATGCTCCAGACACG 576
QY 45 LysValAsnPhenylalanylArgAsn***Ser***GlnGlnAlaValGluValTrp 64
DB 577 AAAGTTAATTTCTATGCTCGGAAGAGATGAGAGTCTGCGGCGCAGGCGCTGCAACTCT 636
QY 65 GlnGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeuLeuValAsnSer 84
DB 637 CAGGCGCTGCGGCTCTGCTGCGAAGCTGCTGCGGCGCAGGCGCTGCTGCAACTCT 696
QY 85 SerGlnProTrpGluProLeuGlnLeuHisValAlaAspLysAlaValSerGlyLeuArgSer 104
DB 697 TCCAGCGCTGAGGAGCCCTGCAAGCTGATGATGAAGCCGTCAAGTCTGCGCAGC 756
QY 105 LeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGluAlaIleSerProProAspAla 124
DB 757 CTCACCACTGCTCTCGGCTCTGAGAGCCCGAAGGAAGCAATCTCCCTCCAGATGCG 816
QY 125 AlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLysLeuPheArgVal 144
DB 817 GCCTCAGCTGCTCCACTCCGAACATCATGCTGACACTTCCGCAACTCTCCGAGTC 876
QY 145 TyrSerAsnPhenylalanylLysLeuLysLeuTyrThrGlyGluAlaCysArgThrGly 164
DB 877 TACTCAATTTCTCTCGGAGAAAGCTGAAGCTGTACACAGGAGGCGCTGACAGCAGAGG 936
QY 165 AspArg 166
DB 937 GACAGA 942

RESULT 9
US-08-750-128-4
Sequence 4, Application US/08750128
Patent No. 5916773
GENERAL INFORMATION:
APPLICANT: MELE Antonio,
APPLICANT: DE SANTIS Rita,
APPLICANT: CARLONI Cristina,
APPLICANT: COSCARIELLA Annamaria
TITLE OF INVENTION: Hybrid molecule of formula GM-CSF-L-EPO or
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: LOWE, PRICE, LeBLANC & BECKER
STREET: 99 Canal Center Plaza, Suite 300
CITY: Alexandria
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk, 3.50 inch.
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/750,128
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IT FI 94 A 000106
FILING DATE: 27-MAY-1994
ATTORNEY/AGENT INFORMATION:

NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER
STREET: 99 Canal Center Plaza, Suite 300
CITY: Alexandria
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk, 3.50 inch.
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/750,128
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IT FI 94 A 000106
FILING DATE: 27-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Robert L. Price
REGISTRATION NUMBER: 22,685
REFERENCE/DOCKET NUMBER: 2879-007
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-684-1111
TELEFAX: 703-684-1124
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 969 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
US-08-750-128-4

Alignment Scores:
Pred. No.: 1,61e-97 Length: 969
Score: 802.00 Matches: 158
Percent Similarity: 97.53% Conservative: 0
Best Local Similarity: 97.53% Mismatches: 4
Query Match: 95.14% Indels: 0
DB: 2 Gaps: 0

US-09-813-775c-18 (1-166) x US-08-750-128-4 (1-969)

QY 5 LeuIleCysAspSerArgValLeuGluArgTyrLeuLeuGluAlaLysGluAlaGlnAsn 24
|||||
DB 481 CTCATCTGTACACGCCAGTCCTGGAGAGGTAACCTTTGGAGGCCCAAGAGCCGCAAGAT 540

QY 25 IleThrThrglyCysAlaGluHisCysSerLeuAsnGluAsnIleThrValProAspThr 44
|||||
DB 541 ATCCAGCAGCGGCGTGTCTGAACACTGCACCTTGAAATATCACTATGCCAGACACC 600

QY 45 LysValAsnPhetYrAlaTrpLysArgAsn***Ser**GlnGlnAlaValGluValTyrP 64
|||||
DB 601 AAAGTTAATTCTATGCTGTGGAAGAGGTGAGGTCCGGCAGCAGCCCGTAAGAAGTCTGG 660

QY 65 GlnGlyLeuValLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeuValAsnSer 84
|||||
DB 661 CAGGGCCCTGGCCCTGCTGTGGAGAGCTGCTCCGCGGCGCAGGCCCTGTGCTCAACTCT 720

QY 85 SerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSerGlyLeuArgSer 104
|||||
DB 721 TCCCAAGCCGTGGAGCCCTGCACCTGCATGTGATAAAGCCGTCACTGGCCTTCGACAC 780

QY 105 LeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGluAlaIleSerProProAspAla 124
|||||
DB 781 CTCACCACTGTGCTTGGGCTGTGGAGCCAGAAAGAACCATCTCCCTCCAGATGCG 840

QY 125 AlaSerIalaIleProLeuArgThrIleThrAlaAspThrIleArgLysLeuPheArgVal 144
|||||

DB 841 GCCCTACCTGCTCCACGTCCGACAAATCATCTCTGACACTTTCCGCAACTCTTCGAGTC 900

QY 145 TyrSerAsnPhetLeuArgLysLysLeuLysLeuTyrThrGlyGluAlaCysArgThrGly 164
|||||

DB 901 TACTCCAAATTCTCCCGGGGAACCTGAAGCTGTACACAGGGAGGCCCTGCAGACAGG 960

QY 165 AspArg 166
|||||

DB 961 GACAGA 966

RESULT 10
US-08-910-647-3
Sequence 3, Application US/08910647
Patent No. 6251433
GENERAL INFORMATION:
APPLICANT: Zuckermann et al.
TITLE OF INVENTION: Compositions and Methods for
TITLE OF INVENTION: Polynucleotide Delivery
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: U.S.A.
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/910,647
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Fujita, Sharon M.
REGISTRATION NUMBER: 38,459
REFERENCE/DOCKET NUMBER: 1218.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 923-2706
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 5107 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-910-647-3

Alignment Scores:
Pred. No.: 1,35e-88 Length: 5107
Score: 743.50 Matches: 148
Percent Similarity: 93.37% Conservative: 7
Best Local Similarity: 89.16% Mismatches: 10
Query Match: 88.20% Indels: 1
DB: 4 Gaps: 1

US-09-813-775c-18 (1-166) x US-08-910-647-3 (1-5107)

QY 1 AlaProProArgLeuIleCysAspSerArgValLeuGluArgTyrLeuLeuGluAlaLys 20
|||||

DB 1791 GCCCAGACAGCCCTCATCTGTGACAGCGAGTCTGGAGAGTACCTCTTGAGAGCCAAAG 1850

QY 21 GluAlaGluAsnIleThrThrglyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
|||||

DB 1851 GAGCCGAGATGTCACAGATGGGCTGTCCGAAAGCTGCAATGCAATATATCACC 1910

QY 41 ValProAspThrLysValAsnPhetYrAlaTrpLysArgAsn***Ser**GlnGlnAla 60
|||||

DB 1911 GTCCAGACACCAAGATTACTCTATGCTGTGGAGAGATGAGAGTGGCGGACAGAGCT 1970

QY 61 ValGluValTrrPcInGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
 DB 1971 GTAGAAAGTCTGGAGGGGCTGGCCCTGCTCTCAGAAAGTGTCTGGGGGAGGCCGCTG 2030
 QY 81 LeuValAsnSerSerGlnProTrrPgluPProLeuGlnLeuHisValAspLysAlaValSer 100
 DB 2031 TTGGCCCAACTCTTCCAGCTTTCGAGCCCTGCGACCTGCACATGATGAACCCATCACT 2090
 QY 101 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGlnAlaLeuSer 120
 DB 2091 GGCCCTTCGAGCATCACACCTGCTGGTGGGGGCTGGAGGCCAG---GAACCATCTCC 2147
 QY 121 ProProAspAlaAlaSerAlaAlaProLeuArgThrThrThrAlaAspThrPheArgLys 140
 DB 2148 CTCACAGATGCGGCTCGGCTGCTCCACTCCGACCATCATCTGACACTTCTGCAAA 2207
 QY 141 LeuPheArgValTyrSerAsnPheLeuArgGlyLysLeuLysLeuThrGlyGlnAla 160
 DB 2208 CTCTTCGAGTCTACTCCCAATTTCCTCCGGGAAAGCTGAAGCTGTACAGGGGAGGCC 2267
 QY 161 CysArgThrGlyAspArg 166
 DB 2268 TGCAGAGAGGGGACAGA 2285

RESULT 11

US-09-620-925-3

Sequence 3, Application US/09620925

Patent No. 6468986

GENERAL INFORMATION:

APPLICANT: Zuckermann et al.

TITLE OF INVENTION: Compositions and Methods for Polynucleotide Delivery

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Chiron Corporation

STREET: 4560 Horton Street

CITY: Emeryville

STATE: California

COUNTRY: U.S.A.

ZIP: 94608-2916

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/620,925

FILING DATE: 21-Jul-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/910,647

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Fujita, Sharon M.

REGISTRATION NUMBER: 38,459

REFERENCE/DOCKET NUMBER: 1218.002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (510) 923-2706

TELEFAX: (510) 655-3542

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 5107 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Alignment Scores:

Pred. No.: 1,35e-88 Length: 5107
 Score: 743.50 Matches: 148
 Percent Similarity: 93.37% Conservative: 7

Best Local Similarity: 89.16% Mismatches: 10
 Query Match: 88.20% Indels: 1
 DB: 4 Gaps: 1

US-09-813-775c-18 (1-166) x US-09-620-925-3 (1-5107)

QY 1 AlaProProArgLeuIleCysAspSerArgValLeuGluArgTyrLeuLeuGlnAlaLys 20
 DB 1791 GCCCCACACGCTCATCTGTACAGCCGAGTCCTGAGAGAGTACTCTTGGAGGCCAAG 1850
 QY 21 GluAlaGlnAsnIleThrThrGlyCysAlaGlnHisCysSerLeuAsnGlnAsnIleThr 40
 DB 1851 GAGGCCGAGAAATGTCAGATGGGCTGTTCCGAAGCTGCAGCTTGAATGAGATATCAC 1910
 QY 41 ValProAspThrLysValAsnPheTyrAlaTrrPlysArgAsn***Ser**GlnGlnAla 60
 DB 1911 GTCCAGACACCAAAAGTTAACTTATGCTGTGAAGAGATGAGAGTGGGAGAGAGCT 1970
 QY 61 ValGluValTrrPcInGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
 DB 1971 GTAGAAAGTCTGGAGGGGCTGGCCCTGCTCTCAGAAAGTGTCTGGGGGAGGCCGCT 2030
 QY 81 LeuValAsnSerSerGlnProTrrPgluPProLeuGlnLeuHisValAspLysAlaValSer 100
 DB 2031 TTGGCCCAACTCTTCCAGCTTTCGAGCCCTGCGACCTGCACATGATGAACCCATCACT 2090
 QY 101 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGlnAlaLeuSer 120
 DB 2091 GGCCCTTCGAGCATCACACCTGCTGGTGGGGGCTGGAGGCCAG---GAACCATCTCC 2147
 QY 121 ProProAspAlaAlaSerAlaAlaProLeuArgThrThrThrAlaAspThrPheArgLys 140
 DB 2148 CTCACAGATGCGGCTCGGCTGCTCCACTCCGACCATCATCTGACACTTCTGCAAA 2207
 QY 141 LeuPheArgValTyrSerAsnPheLeuArgGlyLysLeuLysLeuThrGlyGlnAla 160
 DB 2208 CTCTTCGAGTCTACTCCCAATTTCCTCCGGGAAAGCTGAAGCTGTACAGGGGAGGCC 2267
 QY 161 CysArgThrGlyAspArg 166
 DB 2268 TGCAGAGAGGGGACAGA 2285

RESULT 12

PCT-US94-04361-22

Sequence 22, Application PC/TUS9404361

GENERAL INFORMATION:

APPLICANT: Brigham and Women's Hospital

APPLICANT: 75 Francis Street

APPLICANT: Boston, MA 02115

APPLICANT: Bunn, H. Franklin

APPLICANT: Wen, Danyl

APPLICANT: Showers, Mark O.

TITLE OF INVENTION: Erythropoietin Mutelins With Enhanced

NUMBER OF SEQUENCES: 59

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox

STREET: 1100 New York Avenue, Suite 600

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/04361

FILING DATE: Herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/049,802

```

FILING DATE: 21-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Cimbala, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0627.336PC01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 681 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: both
PCT-US94-04361-22

Alignment Scores:
Pred. No.: 1, 11e-89 Length: 681
Score: 741.50 Matches: 147
Percent Similarity: 93.37% Conservative: 8
Best Local Similarity: 88.55% Mismatches: 10
Query Match: 87.96% Indels: 1
DB: 5 Gaps: 1

US-09-813-775C-18 (1-166) x PCT-US94-04361-22 (1-681)
QY 1 AlaProProArgLeuLeuIleCysAspSerArgValLeuGluArgTyrLeuLeuGluAlaLys 20
DB 72 GCGCCGACGAGCGCTGCTGTGTACAGCCGAGCTCTGGAGAGTACTCTTGGAGGCCAAG 131
QY 21 GluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
DB 132 GAGGCCGAGAAATGTCAGCATGGCGCTGTCGAAAGCTGCAGTGTGAATGAAATATCACC 191
QY 41 ValProAspThrLysValAsnPhetYrAlaTrpLysArgAsn***Ser***GlnGlnAla 60
DB 192 GTCCGACAGACCAAAAGTTAACTTATAGCTGAGAGAGGATAGAGCTCGGCGACAGGCT 251
QY 61 ValGluValTrpGlnGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
DB 252 GTAGAAGTCTGGCAGGCGCTGGCGCTGCTCTCAGAAAGCTGCTGGGGGCCAGGCCGTG 311
QY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSer 100
DB 312 TTGGCCCAACTCTCCAGGCTTTCGAGCCCTGCACCTGACATGATAAAGCCATCAGT 371
QY 101 GlysLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnGlyAlaIleSer 120
DB 372 GGCCTTCGACGATCACCACTGCTGCTGGCGCTGGAGCCAG--GAGGCATCTCC 428
QY 121 ProProAspAlaAlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLys 140
DB 429 CTCGCCAGATGGCGGCTCGGCTGCTCCAGTCGGAACCATCAGCTCTACACTTTCGAAA 488
QY 141 LeuPheArgValItySerAsnPhelLeuArgGlyLysLeuLysLeuTyrThrGlyGluAla 160
DB 489 CTCCTCCGAGTACTACTCAATTTCCTCGGGGAAAGCTGAAGCTTAACACGGGGGAGGCC 548
QY 161 CysArgThrGlyAspArg 166
DB 549 TGCAGGACAGGGGACAGA 566

```

```

TITLE OF INVENTION: Activity
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04361
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/049,802
FILING DATE: 21-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Cimbala, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0627.336PC01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 681 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: both
PCT-US94-04361-27

Alignment Scores:
Pred. No.: 1.79e-83 Length: 681
Score: 695.00 Matches: 137
Percent Similarity: 87.35% Conservative: 8
Best Local Similarity: 82.53% Mismatches: 21
Query Match: 82.44% Indels: 0
DB: 5 Gaps: 0

US-09-813-775C-18 (1-166) x PCT-US94-04361-27 (1-681)
QY 1 AlaProProArgLeuLeuIleCysAspSerArgValLeuGluArgTyrLeuLeuGluAlaLys 20
DB 67 GCGCCGACGAGCGCTGCTGTGTACAGCCGAGCTCTGGAGAGTACTCTTGGAGGCCAAG 126
QY 21 GluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
DB 127 GAGGCCGAAATGTCAGCATGGCGCTGTGTGAAGGCTGCAGCTTCACTGGAATATCACC 186
QY 41 ValProAspThrLysValAsnPhetYrAlaTrpLysArgAsn***Ser***GlnGlnAla 60
DB 187 GTCCGACAGACCAAAAGTTAACTTATAGCTGAGAGAGGATGACAGCTCGGCGACAGGCT 246
QY 61 ValGluValTrpGlnGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
DB 247 GTGGAAGTCTGGCAGGCGCTGGCGCTGCTCTCAGAAAGCTGCTGGGGGCCAGGCCGTG 306
QY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSer 100
DB 307 CTGGCCAACTCTCCAGCATGTGAGACCTCTGACGTGATGATAAAGCCGTCAGC 366
QY 101 GlysLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnGlyAlaIleSer 120
DB 367 AGCCTGCGAGCTCACTCTGCTTGGGCTGCTGGAGCCAGAAAGGACCACTCC 426
QY 121 ProProAspAlaAlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLys 140
DB 427 CTCGACGAGGACCACTCTGCTGCTCCAGCTTCGAAACATTCAGTGTGATCTTGTGAAA 486

```

QY 141 LeupheargValTyrSerAsnPhelenuarglyLysLeuTyrThrGlyAla 160
DB 487 CTTTCCGAATCTACTCCAACTCTCGGGGAGAGCTGACACAGGAGGCC 546
QY 161 CysArgThrGlyAspArg 166
DB 547 TCCGAGAGAGGAGCAGG 564

RESULT 14
PCT-US94-04361-24
Sequence 24, Application PC/TUS9404361
GENERAL INFORMATION:
APPLICANT: Brigham and Women's Hospital
APPLICANT: 75 Francis Street
APPLICANT: Boston, MA 02115
APPLICANT: Bunn, H. Franklin
APPLICANT: Wen, Danyl
APPLICANT: Showers, Mark O.
TITLE OF INVENTION: Erythropoietin Muteins With Enhanced
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04361
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/049,802
FILING DATE: 21-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Cimbal, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0627,336PC01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 678 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: both
PCT-US94-04361-24

Alignment Scores:
Pred. No.: 4,46e-83 Length: 678
Score: 692.00 Matches: 135
Percent Similarity: 88.55% Conservative: 12
Best Local Similarity: 81.33% Mismatches: 19
Query Match: 82.09% Indels: 0
Caps: 0

US-09-813-775c-18 (1-166) x PCT-US94-04361-24 (1-678)

QY 1 AlaProPArgLeuIleCysAspSerArgValLeuGluArgTyrLeuLeuGluAlaLys 20
DB 69 GCTCCGCCACGCGCTATTGGAGAGTCGCGTTCGAGAGGTACATCTTGAGGCCAAG 128
QY 21 GluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
DB 129 GAGGAGAAATGTCAATGAGGCTGTGCAGAAAGTCCCAAGACTGAGAGAAATATTACC 188

QY 41 ValProAspThrLysValAsnPhetYrAlaTrpLysArgAsn**Ser**GlnGlnAla 60
DB 189 GTCCAGATATACAAAGTCATCTTACGCTGGAAAGATGAAAGCTGGAAGACAGGCT 248
QY 61 ValGluValTrrpGlnGlyLeuAlaLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
DB 249 GTAGAAATTTGGCAAGGCGCTCTGCTCTCAGAACCATCTGACGCCAGCGCTGTG 308
QY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSer 100
DB 309 CAGGCCAATTTCTCCACGACCCAGACAGAGCTTACGCTTCATATGACAAAGCCATCAGT 368
QY 101 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGluAlaLysSer 120
DB 369 GGGCTACGAGCGCTCATTCTGCTGCGGCTCGGAGAGCTGCAAGAAATGATGTGCG 428
QY 121 ProProAspAlaAlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLys 140
DB 429 CTTCCAGACGCCCAAGCGCTCCACTCCGACACATCAGCGAGATCTTCTGCAAG 488
QY 141 LeupheargValTyrSerAsnPhelenuarglyLysLeuTyrThrGlyAla 160
DB 489 CTTTCCGGGCTTACTCCAACTTCTCGGGGAGAGACTGTAAGCTGTACACGGGGAGGCC 548
QY 161 CysArgThrGlyAspArg 166
DB 549 TCCAGCAGAGGAGCAGG 566

RESULT 15
PCT-US94-04361-23
Sequence 23, Application PC/TUS9404361
GENERAL INFORMATION:
APPLICANT: Brigham and Women's Hospital
APPLICANT: 75 Francis Street
APPLICANT: Boston, MA 02115
APPLICANT: Bunn, H. Franklin
APPLICANT: Wen, Danyl
APPLICANT: Showers, Mark O.
TITLE OF INVENTION: Erythropoietin Muteins With Enhanced
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04361
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/049,802
FILING DATE: 21-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Cimbal, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0627,336PC01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 679 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPLOGY: both
PCT-US94-04361-23

Alignment Scores:

Pred. NO.:	6, 1le-81	length:	679
Score:	676.00	Matches:	130
Percent Similarity:	86.75%	Conservative:	14
Best Local Similarity:	78.31%	Mismatches:	22
Query Match:	80.19%	Indels:	0
DB:	5	Gaps:	0

US-09-813-775C-18 (1-166) x PCT-US94-04361-23 (1-679)

QY	1	AlAProPoaAgLeuIleCysAspSerArgValLeuLurArgIrrLeuLeuGluAlaIys	20
Db	69	GCTCCCCACCCCTCACTCGAGACAGTCGACGTTCTGGAGAGATCACTTAGACGCCAAC	128
QY	21	GIuAlaGluasnIleThrThrGlyCysAlaGluHisCysSerLeuynGlnIuasnIleThr	40
Db	129	GAGCAGAAATGTGCAGATGGGTGTGGCAGAAAGTCCAGACTGAGTGAATATTACA	188
QY	41	ValProAspThrIlysValAsnPhetYrAlaTriPlysArgsn***Ser***GIgInIuaIa	60
Db	189	GTCCCAGATACCAAGTCACACTTATGCTTGGAAAAGAAATGGAGGTGGAGAAGACAGGCC	248
QY	61	ValGIuValITrpgInglyLeuAlaLeuLeuSerGlnIuaIaValLeuArgGlyGlnAlaLeu	80
Db	249	ATPACAGTTGGCAGGCCGTGCTCCTCTCAGAAACCATCTGCGAGGCCAGGCCCTG	308
QY	81	LeuValAsnSerSerGlnProITrpgIuProLeuGlnIleuHisValAspIlysAlaValSer	100
Db	309	CTACCCAAATTCCTCCAGCCACAGAGAACCTTACGTTATATAGCAAAAGCCATCAGT	368
QY	101	GlyLeuArgSerLeuThrThrLeuLeuArgAlaIleuGlyIaGlnyGlnAlaIaIeser	120
Db	369	GGCTACGATACCTCACTTCACTGCTCGGGTACTGGAGACTCAGAAAGAAATGATGTGC	428
QY	121	ProProAspAlaIaSerAlaIaProLeuArgThrIleThrAlaAspThrPheArgIys	140
Db	429	CTCTCAGATACACCCCACTGCTGCCACTCCGAACATCTACAGTGGATCTTCTCGAAC	488
QY	141	LeuPheArgValIrrSerAsnPhleLeuArgIlyLysLeuIysLeuIrrThrGlyGlnAla	160
Db	489	CTCTCCGGGCTCTACGCGCACTTCCTCCGGGGAAACTGAAGCTGTACAGGGAGAGGTC	548
QY	161	CysArgThrGlyAspArg	166
Db	549	TGCAGAGAGGGGACAGG	566

Search completed: January 8, 2003, 05:36:48
Job time : 44.454 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 8, 2003, 05:18:00 : Search time 42.5404 Seconds
(without alignments)
1717.360 Million cell updates/sec

Title: US-09-813-775c-18

Perfect score: 843
Sequence: 1 APPRLICDSRYLERYLLEAK.....NFLRGLKLYTGECACRTGDR 166

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 389086 seqs, 220051671 residues

Total number of hits satisfying chosen parameters: 778172

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=x1h
-MODEL=frame+ p2n.model -DB=Published Applications NA -QEXT=fastap -SUFFIX=p2n.rmpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORR=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09813775_@CGN_1.1.25 @runat.07012003.153123.23862
-NCPU=6 -ICPU=3 -NO_XLPEXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA:

1: /cgnt2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgnt2_6/ptodata/1/pubpna/PCr_NEW_PUB.seq:*
3: /cgnt2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgnt2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgnt2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgnt2_6/ptodata/1/pubpna/PCrUS_PUBCOMB.seq:*
7: /cgnt2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgnt2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgnt2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgnt2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgnt2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgnt2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgnt2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgnt2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	567	67.3	2426	10	US-09-975-063-1
2	567	67.3	2426	12	US-10-011-858-1
3	325	38.6	423	10	US-09-864-761-16087
4	284	33.7	191	10	US-09-864-761-32584

5	81.5	9.7	2863	9	US-10-136-224-2	Sequence 2, Appl1
6	81.5	9.7	3686	9	US-10-136-224-3	Sequence 3, Appl1
7	81.5	9.7	3940	9	US-10-136-224-1	Sequence 1, Appl1
8	79	9.4	1663	10	US-09-070-927A-277	Sequence 277, App
9	79	9.4	2022	10	US-09-782-980-69	Sequence 67, Appl
10	79	9.4	2815	10	US-09-782-980-67	Sequence 67, Appl
11	76.5	9.1	7419	10	US-09-815-242-4009	Sequence 67, Appl
12	73	8.7	174424	10	US-09-967-768A-314	Sequence 4109, Ap
13	72	8.5	1290	10	US-09-764-847-1605	Sequence 314, App
14	71	8.4	1443	10	US-09-901-884-15	Sequence 1605, Ap
15	71	8.4	2000	10	US-09-887-566-853	Sequence 15, Appl
16	71	8.4	2658	10	US-09-815-242-4035	Sequence 853, App
17	71	8.4	3885	10	US-09-815-242-9626	Sequence 4035, Ap
18	71	8.4	9500	10	US-09-901-884-21	Sequence 9626, Ap
19	70.5	8.4	17286	10	US-09-764-847-3234	Sequence 21, Appl
20	70.5	8.4	1379	10	US-09-974-300-2909	Sequence 3234, Ap
21	70.5	8.4	2332	9	US-09-764-860-136	Sequence 2909, Ap
22	70.5	8.4	2332	10	US-09-764-860-1197	Sequence 136, Appl
23	70	8.3	554	10	US-09-864-761-15332	Sequence 1197, Ap
24	70	8.3	1934	10	US-09-909-325-1	Sequence 15332, A
25	70	8.3	1934	10	US-09-909-325-1	Sequence 1, Appl1
26	70	8.3	1934	10	US-09-909-326-1	Sequence 1, Appl1
27	70	8.3	180557	12	US-10-003-806-6	Sequence 6, Appl1
28	70	8.3	180557	12	US-10-003-806-9	Sequence 9, Appl1
29	69.5	8.2	2322	9	US-10-045-815-7	Sequence 9, Appl1
30	69.5	8.2	2416	9	US-10-045-815-5	Sequence 5, Appl1
31	69.5	8.2	2442	9	US-09-964-899-24	Sequence 24, Appl1
32	69	8.2	1119	10	US-09-974-300-4290	Sequence 4290, Ap
33	69	8.2	13957	10	US-09-782-378A-22	Sequence 22, Appl
34	69	8.2	13957	10	US-09-880-107-2284	Sequence 2284, Ap
35	68.5	8.1	4695	10	US-09-801-368-243	Sequence 243, App
36	68.5	8.1	21423	10	US-09-764-877-2835	Sequence 2835, Ap
37	68.5	8.1	40392	10	US-09-954-456-44	Sequence 44, Appl
38	68.5	8.1	40392	10	US-09-954-456-67	Sequence 67, App
39	68	8.1	1149	9	US-09-938-842A-2167	Sequence 2167, App
40	68	8.1	1171	9	US-09-945-182-25	Sequence 25, Appl
41	68	8.1	2149	10	US-09-925-301-130	Sequence 130, App
42	68	8.1	7626	9	US-10-001-835-82	Sequence 82, Appl
43	68	8.1	8160	10	US-09-070-927A-159	Sequence 159, App
44	67.5	8.0	332	10	US-09-878-574-2932	Sequence 2932, Ap
45	67.5	8.0	4605	10	US-09-070-927A-315	Sequence 315, App

ALIGNMENTS

RESULT 1
US-09-975-063-1
Sequence 1, Application US/09975063
Patent No. US20020045255A1
GENERAL INFORMATION:
APPLICANT: POWELL, Jerry S.
TITLE OF INVENTION: HUMAN ERYTHROPOIETIN GENE: HIGH LEVEL
EXPRESSION IN STABLY TRANSFECTED MAMMALIAN CELLS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dorsey & Whitney, LLP
STREET: Suite 3400, 1420 Fifth Avenue, U.S. Bank Centre
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.7/78
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/975,063
FILING DATE: 10-Oct-2001
ATTORNEY/AGENT INFORMATION:
NAME: Roberts, Mark W.
REGISTRATION NUMBER: 46,160
REFERENCE/DOCKET NUMBER: 500582.03 (112893.109)

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 206-903-8728
 TELEFAX: 206-903-8820
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2426 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-975-063-1

Alignment Scores:

Pred. NO.:	1,02e-63	Length:	2426
Score:	567.00	Matches:	160
Percent Similarity:	32.34%	Conservative:	2
Best Local Similarity:	31.94%	Mismatches:	4
Query Match:	67.26%	Indels:	337
DB:	10	Gaps:	3

US-09-813-775c-18 (1-166) x US-09-975-063-1 (1-2426)

```

OY 1 AAlAPROAArgLeuIleCysAspSerArgValLeuGluArgTyrLeuLengLualys 20
    |||||||
DB 706 GCCCACCACGCTCATCTGTACAGCCGAGCTCTGAGAGGTACCTCTTGGAGGCCAAG 765
OY 21 GluAlaGluAsnIleThr-----
    :|||||
DB 766 AAGCCGGAATATTCAGCGGTGAGACCCTTCCCGACACATTCACAGAACTCAGCTCA 825
OY 26 -----
DB 826 GGGCTTCAGGAACTCTCCAGATCCAGAACTGCACTTGTGGGGTGAGATTGG 885
OY 26 -----
DB 886 GAACTAGACACTGCCCCCTACATAGATATAGTGTGTGGCCCAACCATACCTGGA 945
OY 26 -----
DB 946 AACTAGGCAAGGAGCAAAAGCCAGACGATCTAGCGCTGTGGCCAGGCGCAGAGCTTCAG 1005
OY 27 -----
DB 1006 GGACCTTGACTCCCGGGCTGTGTGCAATTCAGAGCGGCTGTCTGAAACACTGCAGCTT 1065
OY 35 uAsnGluAsnIleThrValProAspThrLysAlaAspThrLysArgAsn-- 54
    |||||||
DB 1066 GAATGAAATATCTACTGTCTCCAGACCAAGATTAATTTCTATGCTTGGAGAGATGGA 1125
OY 54 -----
DB 1126 GGTGAGTCTTTTTTTTTTTTTTTTCTTTTGGAGAAATCTCATTTGGAGAGCTGAT 1185
OY 54 -----
DB 1186 TTTGATGAAAGAGAAATGATCGAGGAAAGTAAATGAGAGCAGAGATGAGCTG 1245
OY 54 -----
DB 1246 CCTGGGCGCAGAGGCTACGCTCTATATCCAGGCTGAGATGGCCGAGATGGGAGAAATTG 1305
OY 54 -----
DB 1306 CTGAGCCCGAGATTTCAGACCAACTAGGAGCATAGTAGATCCCATCTCTACAA 1365
OY 54 -----
DB 1366 ACAATTAATAAATATGACAGGTGAAGTGTGATGTGTAGTCCAGATATTTGGAAG 1425
OY 54 -----
  
```

```

DB 1426 GGTGAGGCGGAGGATCGCTGGAGCCCAAGAAATTTGAGGCTGCACTGTGATCA 1485
OY 54 -----
DB 1486 CCACCTGAACCTCCAGCCTCAGTACAGAGTGAAGGCCCTGTCTCAAAAAAGAAAAA 1545
OY 54 -----
DB 1546 GAAAAATATAGAGCGGTATGAAATAGCTTATTCATTCACTCACTCACTCACTCA 1605
OY 54 -----
DB 1606 TTCATTCAATTCATTCATTCACAAAGCTTATTCATACCTCTCTGTTCAGCTTGGTG 1665
OY 54 -----
DB 1666 CTTGGGCGCTCGTGAAGGCGGAGGAGGAGGATGACATCCCTCAGCTGACCTCCAGATC 1725
OY 55 -----
DB 1726 CACTCCCTGTAGTGTGGG-CAGCAGCGCGTAGAACTGTGGCAGGCGCGCTGCTGTG 1784
OY 71 rGluValValLeuArgGlyGlnAlaLeuValAsnSerSerGlnProTyrPglProle 91
    |||||||
DB 1785 GGAAGCTGTCTGCGGGGCGCAGGCCCTGTGTGTAACCTTCCAGCCGTGGAGCCCT 1844
OY 91 uGlnLeuHisValAspLysAlaValIserGlyLeuArgSerLeuThrThrLeuLeuArgAl 111
    |||||||
DB 1845 GCAGCTGCATGTGATTAAGCGCTCAGTGGCTTGGCAGCCCTCACACTCTCTCGGGC 1904
OY 111 aLeuGlyAla-----
    |||||||
DB 1905 TCTGGGAGC-CCAGGTAGTAGAGAGCAGCACTTCTGCTTGCCTTCTGTAAAGAGGG 1963
OY 114 -----
DB 1964 AGAAGGCTTCTTCTAAGAGATACAGAACTGCTGATTCCTTCCCTTCTGTGGCACTG 2023
OY 115 -----
DB 2024 CAGGAGCTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2083
OY 126 rAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLysLeuPheArgValTyrSe 146
    |||||||
DB 2084 AGCTGCTCCACCTCCGAAACATACAGCTGACACTTCCGCAAACTCTCCGAGTACTC 2143
OY 146 rAsnPheLeuArgGlyLysLeuLysLeuThrGlyGluAlaCysArgThrGlyAspAr 166
    |||||||
DB 2144 CAATTCCTCCGGGAAAGCTGAAGCTGTACACAGGGAGGCGCTTGCAGGACAGGGAGCAG 2203
OY 166 g 166
DB 2204 A 2204
  
```

RESULT 2
 US-10-011-858-1
 Sequence 1, Application US/10011858
 Patent No. US20020137145A1

GENERAL INFORMATION:

APPLICANT: POWELL, Jerry S.
 TITLE OF INVENTION: HUMAN ERYTHROPOIETIN GENE: HIGH LEVEL
 EXPRESSION IN STABLY TRANSFECTED MAMMALIAN CELLS
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESS: Dorsey & Whitney, LLP
 STREET: Suite 3400, 1420 Fifth Avenue, U.S. Bank Centre
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98101
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: MS-WORD
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/011,858
 FILING DATE: 05-NO. US20020137145A1-2001
 ATTORNEY/AGENT INFORMATION:
 NAME: Roderts, Mark W.
 REGISTRATION NUMBER: 46,160
 REFERENCE/DOCKET NUMBER: 500582.13
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 206-903-8728
 TELEFAX: 206-903-8820
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2426 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-10-011-858-1

Alignment Scores:
 Pred. No.: 1,02e-63 Length: 2426
 Score: 567.00 Matches: 160
 Percent Similarity: 32.34% Conservative: 2
 Best Local Similarity: 31.94% Mismatches: 4
 Query Match: 67.26% Indels: 337
 DB: Gaps: 3
 US-09-813-775C-18 (1-166) x US-10-011-858-1 (1-2426)
 QY 1 AAlAPrOArGLeuIleCYsAsPSeArQVAlleuGlaRTyrLeuIleuGlaAlaLys 20
 Db 706 GCGCCACGACGGCTGCTGTCGTGACAGCCAGTCTCTGAGAGTACTCTTGAGGCCAAG 765
 QY 21 GluAlaGluAsnIleThr----- 26
 Db 766 AAGGCCGAGAAATATCAGGGTGAGACCCCTTCCCGACATTCACAGAACTCAGCTCA 825
 QY 26 ----- 26
 Db 826 GGGCTTCACGGCACTCTCTCCAGATCCAGGAACCTGGACCTTGTTGGGGTGGAGTTGG 885
 QY 26 ----- 26
 Db 886 GAAGCTAGACACTGCCCCCTACATAAGAATAAGTGTGGGCCCAAAACCATACCTGGA 945
 QY 26 ----- 26
 Db 946 AACTAGGCAAGAGCAAAAGCCAGACAGATCTACGCTGTGGCCAGGGCCAGAGCTTCAG 1005
 QY 27 ----- 35
 Db 1006 GACACCTTGACGCCCGGGCTGTGTCATTTTCAGACGGGCTGTGTCGAACTGCAAGCTT 1065
 QY 35 uAsnGluAsnIleThrValProAspThrLysValAsnPhetyrAlaTrpLysarQAsn-- 54
 Db 1066 GAATGGAATATCAGTGTCCAGACACCAAAAGTTAATTTCTATGCGCTGGAGAAGAGATGGA 1125
 QY 54 ----- 54
 Db 1126 GGTGAGTCTTTTTTTTTTTTTTTCCTTTCTTTTGGAGAAATCATTTGCGAGCCTGAT 1185
 QY 54 ----- 54
 Db 1186 TTTTGATCAAAAGGAGATGATCGAGGAAAGTTAAATGAGACGACAGATGAGCCTGG 1245
 QY 54 ----- 54
 Db 1246 CTTGGGCGAGAGGCTCAGCTGTATTAATCCAGAGCTGAGATGGCGAGATGGAGAATTG 1305
 QY 54 ----- 54

Db 1306 CTTGAGCCCGGAGTTTTCAGACCACCACTAGGACGATAGTGAATCCCATCTCTACAA 1365
 QY 54 ----- 54
 Db 1366 ACATTTAAAAAAATTAGTCAGGTGAAGTGTGCATGTGTGTAGTCCACATATTGGGAAG 1425
 QY 54 ----- 54
 Db 1426 GCTGAGGCGGAGAGATCGGTGAGGCCAGCAAGATTGAGGTGACGTGATGATCACA 1485
 QY 54 ----- 54
 Db 1486 CCACTGAACTCAGCCTCAGTACAGAGAGAGGACCTGTCTCAAAAAAGAAAAA 1545
 QY 54 ----- 54
 Db 1546 GAAAAATATAGAGGCTGTATGGAATACGTTCAATTCATTCACCTCACTCACTCA 1605
 QY 54 ----- 54
 Db 1606 TTCATTCAATTCATTCACCAAGCTTATTCATACCTTCTGTGTTGCTCAGCTTGGT 1665
 QY 54 ----- 54
 Db 1666 CTTGGGGCTGTGAGGGGAGAGGAGAGAGGAGACATCCCTCAGCTGACTCCAGAGTTC 1725
 QY 55 -----***Ser***GlnGlnAlaValGluValITPpGlnGlyLeuAlaLeuLeu 71
 Db 1726 CACTCCCTGTAGGTGCGGG-CAGCAGGCGCTAGAAAGTGTGACAGGGCTGGCCCTGCTGTC 1784
 QY 71 rGluAlaValleuAArgIyGlnAlaLeuLeuValAsnSerSerGlnProTrpGluProle 91
 Db 1785 GGAAGCTGTCTCTCGGGGCGAGGCCCTGTGTGGAACTCTTCCAGCCGCGGAGGCCCT 1844
 QY 91 uGlnLeuHisValAspLysAlaValSerGlyLeuAArgSerLeuThrThrLeuLeuAla 111
 Db 1845 GCAGCTGCATGTGTGAATAAGCCGTGAGTGGCTCCAGCTCACCACATCTCTCTGGGG 1904
 QY 111 AluGlyAla----- 114
 Db 1905 TCTGGGAGC-CCAGGTGAGTAGAGCGGACACTTCTGCTTGGCCCTTTCGTAAAGAGCG 1963
 QY 114 ----- 114
 Db 1964 AGAAGGCTCTGCTAAGAGTACAGGAAGTCTCCGTAATCTCTCTTCTGTGGCAGT 2023
 QY 115 -----GlnLysGluAlaIleSerProProAspAlaIase 126
 Db 2024 CAGCGACCTTCCTGTTCTCTCTTGGCAGAAAGCAAGCCATCTCCCTCCAGATCGGCGTTC 2083
 QY 126 rAlaAlaProLeuAArgThrIleThrAlaAspThrPheAArgLysLeuPheAArgValIyrse 146
 Db 2084 AGCTGCTCCACTCCGAACAATCACTGCTGACACTTTCGCAAACTCTTCCGAGTCTACCT 2143
 QY 146 rAsnPheLeuAArgLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLys 166
 Db 2144 CAATTTCTCTCGGGGAAAGCTGAAGCTGTACACAGGGAGAGCCTGCAGAGACAGGACAG 2203
 QY 166 g 166
 Db 2204 A 2204
 RESULT 3
 US-09-864-761-16087/c
 Sequence 16087, Application US/09864761
 Patent No. US20020048763A1
 GENERAL INFORMATION:
 APPLICANT: Penn, Sharon G.
 APPLICANT: Rank, David R.
 APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wensheng
 TITLE OF INVENTION: HUMAN GENOME- DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO

```

: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
: FILE REFERENCE: Aecm1ca-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 16087
: LENGTH: 423
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AF053356.1
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.93
US-09-864-761-16087

Alignment Scores:
Pred. No.: 1,96e-33 Length: 423
Score: 325.00 Matches: 77
Percent Similarity: 62.60% Conservative: 0
Best Local Similarity: 62.60% Mismatches: 1
Query Match: 38.55% Indels: 46
DB: 10 Gaps: 1

US-09-813-775c-18 (1-166) x US-09-864-761-16087 (1-423)
OY 58 GlnGlnAlaValAlaGluValTTPGlnGlyLeuAlaLeuLeuSerGlnAlaValLeuArgGly 77
Db 370 CAGCGGCGGAGAGAGAGTGTGGAGGCGCTGGCTGTGTGGAGAGCTGCTCGCGGGGC 311
OY 78 GlnAlaLeuValAlaSerSerGlnProTPGlnProLeuGlnLeuHisValaPlays 97
Db 310 CAGGCGCTGTGGTCACTCTTCCAGCGCGTGGAGCGCCCTGCACCTCATGTGGATAAA 251
OY 98 AlaValSerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAla----- 114
Db 250 GCCGTCAATGGCCTTCGACAGCTCACCACCTCATCTGCTCGGGCTCTGGAGAC-CCAGGTGAG 192
```

```

OY 114 ----- 114
Db 191 TAGAGCGGACACTTCTGCTTCCCTTCTGTAGAGAGGAGAGGCTTGTAGGA 132
OY 114 ----- 114
Db 131 GTACAGAACTGTCCGTAATTCCTTCTTGTGGCACTGCAGCGACCTCCTGTTTCT 72
OY 115 -----GlnTysGluAlaAlaIleSerProPsaPalaAlaSerAlaAlaProLeuArgThr 132
Db 71 CCTTGCGAAGAGGAGGACATCTCCCTCCAGATGGCGCTAGCTCTCCATCGGAGACA 12
OY 133 lIethAla 135
Db 11 ATCACTGCT 3

RESULT 4
US-09-864-761-32584/C
: Sequence 32584, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
: FILE REFERENCE: Aecm1ca-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 32584
: LENGTH: 191
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
```


Db 526 GAATTAGTCCCTAAGTCTCTGGAACATATGTTTGGCTTAACCAACGAAGTA----- 579
Qy 49 TYRATATPRLYSARGASN***Ser***GlnGlnAlaValAlaGluValTyrGlnGlyLeuAla 68
Db 580 -----CTTCTGGAAACAGTGAAGAAAGCGTCTCAATTGATGAGATTAAGTACTT 627
Qy 69 LeuLeuSerGluAlaValLeuArgGlyGlnAlaLeuLeuValAsnSerSerGlnProTyr 88
Db 628 TTACACAGCAGAGAAACCACTTAAGTTAAATGCAATCTTATGTAATCAG----- 675
Qy 89 GluProLeuGlnLeuHisValAspLysAlaValSerGlyLeuArgSerLeuThrThrLeu 108
Db 676 -----ATTGCGCTTTTCTGCA-----ACAAAAATT 702
Qy 109 LeuArgAlaLeuGlyAlaGlnLysGluAlaIleSerProProAspAlaAlaSerAlaAla 128
Db 703 TTGGCTTAACATGGCGTATCAACACGCAATATTGCT-----CATACATTAAAGGTTTCAT 756
Qy 129 Pro-----LeuArgThrIleThrAlaAspThrPheArgLysLeuPhe 142
Db 757 CCTTATCGAGTCAGTACGCTTACAAACAGTCGACGATTTGAAATGCTCTCGATTAGAA 816
Qy 143 ArgValTyrSerAsnProLeuArgGlyLysLeuLysLeuTyrThrGlyGlu 159
Db 817 AGACTTTATGATGATTAAGTTAGTGAATAATGATTACCGCATGAAACGGGCCAA 867

RESULT 9

US-09-782-980-69
; Sequence 69, Application US/09782980
; Patent No. US20020072089A1
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran M.
; APPLICANT: Macbeth, Kyle J.
; APPLICANT: Busfield, Samantha J.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Gu, Wei
; APPLICANT: White, David
; TITLE OF INVENTION: NOVEL ITALY, LOR-2, STRIFE, TRASH, BDSF, LRSG, AND
; TITLE OF INVENTION: STIMST PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-121CP
; CURRENT APPLICATION NUMBER: US/09/782,980
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: PCT/US00/02125
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 09/448,076
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: 09/276,400
; PRIOR FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: 60/117,580
; PRIOR FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: 09/014,195
; PRIOR FILING DATE: 1998-01-27
; PRIOR APPLICATION NUMBER: 09/014,348
; PRIOR FILING DATE: 1998-01-27
; PRIOR APPLICATION NUMBER: 09/086,892
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 09/296,208
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 09/063,950
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 09/561,381
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/561,810
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/087,121
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 09/672,721
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 09/049,799
; PRIOR FILING DATE: 1998-03-27

; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 69
; LENGTH: 2022
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2019)
US-09-782-980-69

Alignment Scores:
Pred. No.: 1.59 Length: 2022
Score: 79.00 Matches: 32
Percent Similarity: 40.83% Conservative: 17
Best Local Similarity: 26.67% Mismatches: 65
Query Match: 9.37% Indels: 6
DB: 10 Gaps: 3

US-09-813-775c-18 (1-166) x US-09-782-980-69 (1-2022)

Qy 19 AlAlysgLuAlaGlnAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsn 38
Db 458 AGCTCAAGCTGCCAGACATGAGCTTGGGTGTGCCCCATTGCACTTGGCCGCTGC 517
Qy 39 IleThrValProAspThrLysValAsnPheTyrAlaTyrLys-ArgAsn***Ser***G1 58
Db 518 TGCTGCTGACCTCAGCCACACAGCATCCACCCCTCGAAGCCGGAATCTGATACCG 577
Qy 58 ngIn-----AlaValGluVal---TrpGln-GlyLeuAlaLeuSerGluAlaVal 75
Db 578 CCAATGTAGAGCATGATGCTGGCTGAGGCTGAGGCTGCGCGCATGATGAGGG--C 634
Qy 75 euArgGlyGlnAlaLeuLeuValAsnSerSerGlnProTyrGlnProLeuGlnLeuHis 95
Db 635 TTTTGGCGCGCTTCTCAACCTCATGACTTGAATGTTCTTGACAAACCACTTGAGCAT 694
Qy 95 alAspLysAlaValSerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAla 115
Db 695 TGCCATCTGATTCAGAGCGCTGCGCTGACACGCCGCGGCGTGGCAACACCC 754
Qy 115 lAlYsgLuAlaIleSerProProAspAlaAlaSerAlaAlaProLeuArgThrIle 133
Db 755 GTATTGCCAGATACGCGCCGAGGAGCTGCTGTGACTGCTGACCTTACAGAAATTG 810

RESULT 10

US-09-782-980-67
; Sequence 67, Application US/09782980
; Patent No. US20020072089A1
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran M.
; APPLICANT: Macbeth, Kyle J.
; APPLICANT: Busfield, Samantha J.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Gu, Wei
; APPLICANT: White, David
; TITLE OF INVENTION: NOVEL ITALY, LOR-2, STRIFE, TRASH, BDSF, LRSG, AND
; TITLE OF INVENTION: STIMST PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-121CP
; CURRENT APPLICATION NUMBER: US/09/782,980
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: PCT/US00/02125
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 09/448,076
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: 09/276,400
; PRIOR FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: 60/117,580
; PRIOR FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: 09/014,195

PRIOR FILING DATE: 1998-01-27
PRIOR APPLICATION NUMBER: 09/014,348
PRIOR FILING DATE: 1998-01-27
PRIOR APPLICATION NUMBER: 09/086,892
PRIOR FILING DATE: 1998-05-29
PRIOR APPLICATION NUMBER: 09/296,208
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 09/063,950
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 09/561,381
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 09/561,810
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 09/087,121
PRIOR FILING DATE: 1998-05-29
PRIOR APPLICATION NUMBER: 09/672,721
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: 09/049,799
PRIOR FILING DATE: 1998-03-27
NUMBER OF SEQ ID NOS: 176
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 67
LENGTH: 2815
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (197)..(2215)
US-09-782-980-67

Alignment Scores:
Pred. No.: 2.59 Length: 2815
Score: 79.00 Matches: 32
Percent Similarity: 40.83% Conservative: 17
Best Local Similarity: 26.67% Mismatches: 65
Query Match: 9.37% Indels: 6
Gaps: 3

US-09-813-775c-18 (1-166) x US-09-782-980-67 (1-2815)

OY 19 AlAluGluAlaGluAlaSerThrGlyCysAlaGluHisCysSerLeuAsnGln 38
Db 654 AGCTAGAGTGGCAGACATGAGCTTGGGTGCTGCCCCATGACACTGCCCCCTGC 713
OY 39 lIeThrValProAspThrLysValAsnPhetYrAlaTrpLys-ArgAsn**Ser**G1 58
Db 714 TGCTCTTGACCTGACGACACAGACATCCAGCCCTGGAAACCGGAATATCTGATACCG 773
OY 58 nGln-----AlaValGluVal---TrpGln-GlyLeuAlaLeuLeuSerGluAlaValL 75
Db 774 CCAATGTAGAGGCAATGAGGTGGCTGGCTGAGGCTGCGGCAAGCTGATGAGGGC---C 830
OY 75 eUArgGluAlaLeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisV 95
Db 831 TTTTGGCCGCTTCTCAACCTCCATGACTTGATGTTTTCACAAACAGTTGGAGCAT 890
OY 95 aLAspLysAlaValSerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaG 115
Db 891 TGCCATCTGTGATTCAGGCGCTGCGTGGCTGACACAGCGCTGCGGCTGGCAACACCC 950
OY 115 lInLysGluAlaLeuSerProAspAlaAlaSerAlaAlaProLeuArgThrIle 133
Db 951 GTATTGCCAGATACGCGCCGAGGACCTCGTGTCTGACTGCTTACAGGAATTTG 1006

RESULT 11

US-09-815-242-4009
Sequence 4009, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyckind, Judith W.
APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 4009
LENGTH: 7419
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-815-242-4009

Alignment Scores:
Pred. No.: 22.6 Length: 7419
Score: 76.50 Matches: 38
Percent Similarity: 42.00% Conservative: 25
Best Local Similarity: 25.33% Mismatches: 60
Query Match: 9.07% Indels: 27
Gaps: 7

US-09-813-775c-18 (1-166) x US-09-815-242-4009 (1-7419)

OY 8 AspSerArgValLeuGluArgTrpLeuGluAlaLys-----GluAlaGlu 23
Db 3547 GATCCGTTGATGATGAGCGCTTCTCTCGAAGCATGACATCTGCTGAGCCGAG 3606
OY 24 AsnIleThrGlyCysAlaGluHis-----CysSerLeuAsn 36
Db 3607 GACCTGCTGAGCGCTGGCAGACATCCGAGAGCGCCAGAACTGATGCTGCGC 3666
OY 37 GluAsnIleThrValProAspThrLysValAsnPhetYrAlaTrpLysArgAsn**Ser 56
Db 3667 GAGGAGCTCAGACACTCTGAGCGCGCGCC-----CGCATCCCGAG 3708
OY 57 **GlnGlnAlaValGluValTrpGlnGlyLeuAlaLeuLeuSerGluAlaValLeuArg 76
Db 3709 CTGCCACAGGTGGAAGAACTTGCACAGGCTTGTGAGGCTCTAGAGCTCCGAGGAA 3768
OY 77 GlyGln-----AlaLeuValAsnSerSerGlnProTrpGluProLeu 91
Db 3769 GGCCTGCTGGCGGTGAGCCCGGCTTCTTCAGAGAAAGCCGACAGCCGAGCGCTG 3828
OY 92 GlnLeuHisValAspLysAlaValSerGlyLeuArgSerLeuThrThrLeuLeuArgAla 111
Db 3829 ATCGGATGATGAGCAGAGTGGCGCGCGCTGAG---GTACGCGCGCGTCCGAGCGG 3885
OY 112 LeuGlyAlaGlnLysGluAlaLeuSerProAspAlaAlaSerAlaAlaProLeu--- 130
Db 3886 GTCCGCGCGCTGACAGACTGTTGGAAAGCGCC---GCCGCGAGGCGCTTCCCTTATC 3942
OY 131 -----ArgThrIleThrAlaAspThrPhe 138
Db 3943 GACCCGAGAGGCTCTGGCGCGAGCATTTTC 3972

RESULT 12
 US-09-967-768A-314/C
 ; Sequence 314, Application US/09967768A
 ; Patent No. US20020150877A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Augustus, Meena
 ; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
 ; FILE REFERENCE: 689290-72
 ; CURRENT APPLICATION NUMBER: US/09/967,768A
 ; PRIOR FILING DATE: 2001-09-28
 ; PRIOR APPLICATION NUMBER: US/60/236,109
 ; PRIOR FILING DATE: 2000-09-28
 ; PRIOR APPLICATION NUMBER: US/60/236,034
 ; PRIOR FILING DATE: 2000-09-28
 ; PRIOR APPLICATION NUMBER: US/60/236,111
 ; NUMBER OF SEQ ID NOS: 325
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 314
 ; LENGTH: 174424
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-967-768A-314

Alignment Scores:
 Pred. No.: 6.59e+03 Length: 174424
 Score: 73.00 Matches: 39
 Percent Similarity: 34.56% Conservative: 8
 Best Local Similarity: 28.68% Mismatches: 33
 Query Match: 8.66% Indels: 56
 DB: 10 Gaps: 9
 US-09-813-775C-18 (1-166) x US-09-967-768A-314 (1-174424)

QY 3 ProArgLeuIleGysAspSerArgValLeuGluArgTyrLeuLeuGluAlaLysGluAla 22
 |||||
 Db 139850 CCAGGCTGAGTGCAGTATGCCATCTATTACAACTG----- 139812
 QY 23 GluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsn----- 36
 |||||
 Db 139811 -----TGT-----CACTGCAAGCTGATGTCCTGGGTTCAGTGAT 139776
 QY 37 -----GluAsnIleThrValProAspThrLysVal 46
 |||||
 Db 139775 CCTCCCACTCAGCCTCCTGAGTACTGAGTACAAATACACACCATCAAGCCAGCT 139716
 QY 47 Asn-----PheTyrAlaTrpLysArgAsn***Ser**GlnGlnAlaValGluValTrpGln 65
 |||||
 Db 139715 AATTTTATTATTATTGTAAGAA-----TGG--- 139689
 QY 66 GlyLeuAlaIleLeuSerGluAlaValLeuArg-----GlyGlnAlaLeuLeuValAsn 83
 |||||
 Db 139688 GGTCTCCCTACAGTGTCTGCAAGGTGGTCTTAACCTCTGAGGTCAAGCAATCTT----- 139635
 QY 84 SerSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSerGlyLeuArg 103
 |||||
 Db 139634 CTTTCACACCTCCCAAGAGTGTGGATGGACAGCATCAAGCCACTGCGCTGACCTAAAA 139575
 QY 104 -----SerLeuThrThrLeuLeuArg 110
 |||||
 Db 139574 CTAATTTTGTGTGTTTGTGATGAGATGAGTCTACTCTGTGTCGCCAGG 139527

RESULT 13
 US-09-764-847-1605
 ; Sequence 1605, Application US/09764847
 ; Patent No. US20020132767A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PC009
 ; CURRENT APPLICATION NUMBER: US/09/764,847
 ; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 2003
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1605
 ; LENGTH: 1290
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-764-847-1605

Alignment Scores:
 Pred. No.: 6.67 Length: 1290
 Score: 72.00 Matches: 47
 Percent Similarity: 35.35% Conservative: 23
 Best Local Similarity: 23.74% Mismatches: 66
 Query Match: 8.54% Indels: 62
 DB: 10 Gaps: 5
 US-09-813-775C-18 (1-166) x US-09-764-847-1605 (1-1290)

QY 14 ArgTyrLeuLeuGluAlaLysGluAlaGluAsnIleThrThrGlyCysAlaGluHisCys 33
 |||||
 Db 560 CGGTATATCCACAAGCCGTGTAGCAGCGCACATTAAAGCTGTGTCACGCCATTGTGTC 619
 QY 34 SerLeu----- 35
 |||||
 Db 620 AGTCTCTGCGCTCTCAAGCTCAACCCCTCCGCGCACAGGCTCCGACAGCCCAT 679
 QY 36 -----AsnGluAsnIleThrValProAspThrLysValAsnPheTyr 49
 |||||
 Db 680 GCCTCTGACAGCCCTCTCTCGGAGCGCTCCAAAGCCCTGACAGCCGTCTCTCTCG 739
 QY 50 AlaTrp-LysArgAsn***Ser**GlnGlnAlaValGluValTrpGlnLysLeuAla 69
 |||||
 Db 740 TCTTGAAAGGTACAGAGTCTGCCAGCCTATGCTGACAGTGTGGGGCCCTGAACTCT 799
 QY 69 uLeuSerGlu-AlaValLeu-----ArgG 77
 |||||
 Db 800 TTGTATGACAGCGGTCTCTCTGACCTGTGCGCTAGAGAGTGGGAGCGGCGTGGCGG 859
 QY 77 GlnAlaLeuLeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAsp 97
 |||||
 Db 860 GTCATGCAAGGCTCCACAGCCACCTGCGGCGGCTCCAGCAGCAGCAGCAGCAGCAGC 919
 QY 97 ysaAlaValSerGlyLeu-----ArgSerLeuThr 107
 |||||
 Db 920 GCCGAGCGCTGGGCTGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 979
 QY 107 hTrpLeuArgAlaLeuGlyAlaGlnLysGluAlaIleSerProProAspAlaIleSer 127
 |||||
 Db 980 CCTGCTCTGCGCTCTCGGCGTCCAGAGAGCGGTAAGTCTCGGCGGCGGCGGCGGCGG 1039
 QY 127 IaAlaProLeuArgThrIleThrAlaAspThrPheArgLysLeuPhe----- 142
 |||||
 Db 1040 AACCTCTCTTCTTACATTATTAAGT-----TGTCTTCAATTATTTT 1084
 QY 143 -----ArgValTyrSerAsnPheLeuArgGlyLysLeuLys 154
 |||||
 Db 1085 CTTAGGACATTTGATTAATCAATCTTGGATTTTGAGATTAAATAAATTAAA 1134

RESULT 14
 US-09-901-884-15/C
 ; Sequence 15, Application US/09901884
 ; Patent No. US20020037573A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KYOMA HAKKO KOGYO CO., LTD.
 ; TITLE OF INVENTION: F0P1-ATPase polypeptides and their genes
 ; FILE REFERENCE: 113290S1
 ; CURRENT APPLICATION NUMBER: US/09/901,884
 ; CURRENT FILING DATE: 2001-07-09
 ; PRIOR APPLICATION NUMBER: JP 2000-234317
 ; PRIOR FILING DATE: 2000-08-02
 ; NUMBER OF SEQ ID NOS: 21
 ; SOFTWARE: PatentIn version 2.1

```
;; SEQ ID NO 15
;; LENGTH: 1443
;; TYPE: DNA
;; ORGANISM: Corynebacterium ammoniagenes
US-09-901-884-15

Alignment Scores:
Pred. No.: 10.6 Length: 1443
Score: 71.00 Matches: 30
Percent Similarity: 34.62% Conservative: 15
Best Local Similarity: 23.08% Mismatches: 41
Query Match: 8.42% Indels: 44
DB: 10 Gaps: 5

US-09-813-775c-18 (1-166) x US-09-901-884-15 (1-1443)

QY 5 LeuileCysAspSerArgValleuGluArgTyrLeuLeuGluAlaIysGluAlaGluAsn 24
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1284 CTTCTCTGCGAGAGAAAGTTCTG-----GCCCAAGAAAGCGCTCAAT 1243

QY 25 IleThrThrGlyCysAla-----GluHisCysSer 30
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1242 GCGAGC-----TGGCGCGTGAAGCGTGATCTTGTCTCTTCAGACAGCTGCTCATACC 1189

QY 31 -----GluHisCysSer 34
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1188 CAGAAATCCGATATATCTCGCATCTTCTTCTTCGCAAAATGTTGATCAGCGCTG 1129

QY 35 LeuAsnGluAsnIleThrValProAspThrIysValAsnPhenylAlaTryptIysArgAsn 54
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1128 AGCAACACGCGTAGACCGCTCAGCAGCAT-----GCTTGGCTCAGAAAT 1084

QY 55 *****GlnGlnAlaValAluValTyrGlnGlyLeuAlaLeuLeuSerGluAla 73
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1083 ACGAGACGTCGACGACGAGTGGGTTCTGCTGAGATATACCTTGGAAGCAATCGCAGC 1024

QY 74 ValLeuArgGlyGlnAlaLeuLeuValAsnSerSerGlnProTyrGluProLeuGlnLeu 93
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1023 GTCAAGCTCGGGTGTGATCAAGTGAAGCAAGTGCTGCTGAGAGCTGAGTGA 964

QY 94 HisValAspLysAlaValSerGlyLeuArg 103
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 963 GTCATCGGACGAGAGCTAAACGCGCTGCAG 934

RESULT 15
US-09-887-576-853
; Sequence 853, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 853
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-887-576-853
```

```
Alignment Scores:
Pred. No.: 17.1 Length: 2000
Score: 71.00 Matches: 32
Percent Similarity: 47.41% Conservative: 23
Best Local Similarity: 27.59% Mismatches: 41
Query Match: 8.42% Indels: 20
DB: 10 Gaps: 7

US-09-813-775c-18 (1-166) x US-09-887-576-853 (1-2000)

QY 32 HisCys-----SerLeuAsnGluAsnIleThrValProAspThrIysVal 46
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 408 CATTCCTATTCATCTTATCATAGCATTCACCAACACTTCTCATTTCTCCAAAGCTA 467

QY 47 AsnPhenylAlaTryptIysArgAsn***Ser**GlnGlnAlaValGluValTyrGlnGly 66
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 468 AATTTTAGCCATTATATGTAATAAATATGCTCCCAACACATTTCCAACTGCGCAAAAC 527

QY 67 LeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeuLeuValAsnSerSerGln 86
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 528 CTTCTCCACTAGCCAAATTT-----GGTACGCCAAAGTC-----ACTAGCGAA 572

QY 87 ProTyrGluProLeuGlnLeuHisValAspLysAlaValSerGlyLeuArgSerLeuThr 106
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 573 CCGTGG-----GCCCTACACACATTCCTCCCATCTCTCTCGG----- 614

QY 107 ThrLeuArgAlaLeuGlyAla-----GlnLysGluAlaIleSerProProAsp--- 123
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 615 ---CTTCCTCTCGGATGGGCTCGACACACAGCTTCTCCAAAGCTCTGCGCGCTT 671

QY 124 AlaAlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArg 139
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 672 GCCGCGACCTCCACCGCTCTCCACACCGTAGCCACACAGCCGCGC 719

Search completed: January 8, 2003, 07:53:24
Job time : 91.5404 secs
```

TITLE Direct Submission
JOURNAL Submitted (12-APR-2000)

COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES

source
1..1035
location/Qualifiers
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="224N04"
/clone_lib="G"
/note="Genoscope sequence ID : C0AG224DG02SP1-end : PUC-ori"

BASE COUNT 164 a 306 c 302 g 252 t 11 others
ORIGIN

Alignment Scores:

Pred. No.:	11.9	Length:	1035
Score:	82.50	Matches:	34
Percent Similarity:	46.79%	Conservative:	17
Best Local Similarity:	31.19%	Mismatches:	43
Query Match:	9.79%	Indels:	15
DB:	17	Gaps:	5

US-09-813-775C-18 (1-166) x CNS02175 (1-1035)

```

QY 60 AlAValGluValTrpGlnGlyLeuAlaLeuSerGluAlaValLeuArgGlyGlnAla 79
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 227 GCTGCTCCGCATGATACGAGAGAGTGTGCACCTCTCTCTCTCTCCACAGTGTACC 286
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 80 -----LeuLeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspLys 97
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 287 ACCTTCTTATCTGCTCCCTACGATGCTCTG-----GTGGAGATC 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 98 AlAValSerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGlu 117
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 329 GCAGGACAGAGGACTCGCCAGCGCTACAGTACGCGAGCGCTGTCGCCACCCGGGGG 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 118 AlaIleSerProProAspAlaIleSerAlaAlaProLeuArgThrIleThrAlaAspThr 137
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 389 GCCCGCAGCGGCCGCCAGACGCTCTCTCGCTCTCTCTCCAGGATGAGCGCGAGACG 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 138 PheArgLysLeuPheArgValTyrSerAsnPheLeuArgGlyLys---LeuLysLeuTyr 156
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 449 -----CTGGCGATGCGCGGCGCTCTCAGGTCGAGTGTGGCGGTGTAC 499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 157 ThrGly-----GluAlaCysArg 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 500 GACGAGACAGCAGCAGCGGTGTGCG 526
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: January 8, 2003, 07:51:02
Job time : 1407.74 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 8, 2003, 05:16:34 ; Search time 1392.74 Seconds
(without alignments) 1930.339 Million cell updates/sec

Title: US-09-813-775c-18

Perfect score: 843

Sequence: 1 APPRLICDSRVLEERYLEAK.....NFLRGKLYGTGACRTGDR 166

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O/cqnt.1/USPTO.spool/US09813775/runat.07012003.153122.23837/app.query.fasta_1.718
-DB=EST -OPMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdt -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORF=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09813775.@CEN1_1_821.@runat.07012003.153122.23837 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-MARN_TIMEOUT=30 -THRFADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELPEXT=6 -DELEXT=7

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inu:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	441.5	52.4	567	9	AA662379	AA662379 nu93601.s
2	179	21.2	686	13	BM438685	BM438685 IPLPRT0146
3	141	16.7	444	10	BE663803	BE663803 147806 MA
4	120	14.2	968	17	CNS02P50	AL194613 Tetradon
5	106	12.6	365	13	BM279790	BM279790 zah5505 Z
6	105.5	12.5	611	13	BM315886	BM315886 fw66c09.Y
7	92	10.9	593	14	BM931504	BM931504 UI-E-EJ1-
8	92	10.9	657	14	BM185758	BM185758 UI-E-EJ1-
9	87.5	10.4	598	17	A2413715	A2413715 1M0197C23
C 10	86.5	10.3	721	17	A2981594	A2981594 2M0262N06
11	84.5	10.0	759	13	BM072805	BM072805 MEST53-H0
12	84	10.0	2793	11	AK012169	AK012169 Mus muscu
13	83.5	9.9	619	14	BM931641	BM931641 UI-E-EJ1-
14	82.5	9.8	656	17	A2572463	A2572463 306PVC10
15	82.5	9.8	1035	17	CNS02175	AL176522 Tetradon
C 16	81.5	9.7	800	13	BI909650	BI909650 60306680
17	81.5	9.7	899	12	BG173674	BG173674 602336813
C 18	81.5	9.7	908	17	CNS076E5	AL431283 T7 end of
19	81	9.6	867	12	BF184120	BF184120 601843157
20	80.5	9.5	662	13	BM103178	BM103178 fV41h09.Y
C 21	80.5	9.5	1142	14	BM906657	BM906657 AGENCOURT
22	79.5	9.4	524	13	BM437324	BM437324 VVA017G09
23	79.5	9.4	533	14	BQ462002	BQ462002 HD02J17-
24	79.5	9.4	655	10	AV914315	AV914315 AV914315
25	79.5	9.4	675	17	BM466321	BM466321 BOGHX18TF
26	79	9.4	520	13	BF190151	BF190151 BJ490151
27	79	9.4	770	14	BO751539	BO751539 EST632102
28	79	9.4	809	14	BO750964	BO750964 EST631527
29	79	9.4	938	14	BO919901	BO919901 AGENCOURT
C 30	78.5	9.3	895	12	BE914444	BE914444 601668826
C 31	78	9.3	395	10	AV696100	AV696100 AV696100
C 32	78	9.3	584	9	AA198958	AA198958 mU16806.r
C 33	78	9.3	612	10	AV734270	AV734270 AV734270
34	78	9.3	753	12	BF234939	BF234939 602027383
35	78	9.3	832	12	BF310420	BF310420 601895074
36	78	9.3	1002	12	BE963020	BE963020 601656203
37	77.5	9.2	525	14	BO617331	BO617331 fab03a08.
38	77.5	9.2	700	12	BS908975	BS908975 Talc117XG
39	77.5	9.2	804	13	BI088996	BI088996 602835343
40	77.5	9.2	895	10	BE559607	BE559607 601347295
41	77.5	9.2	1375	13	BM542459	BM542459 AGENCOURT
C 42	77	9.1	559	12	BE754300	BE754300 207777 MA
C 43	77	9.1	622	10	AW757556	AW757556 874002A10
C 44	77	9.1	657	14	BP017930	BP017930 BP017930
45	77	9.1	663	12	BG859057	BG859057 1024061B0

ALIGNMENTS

RESULT 1
AA662379/c
LOCUS
DEFINITION
nu93601.s1 NCI_CGAP_Pt22 Homo sapiens CDNA clone IMAGE:1218288 3'
similar to gb:565458 ERYTHROPOIETIN PRECURSOR (HUMAN).; mRNA
sequence.
ACCESSION
AA662379
VERSION
AA662379.1 GI:2616470
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 567)
AUTHORS
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	SOURCE	COMMENT
LOCUS	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
DEFINITION	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
ACCESSION	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
VERSION	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
KEYWORDS	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
ORGANISM	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
SOURCE	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
COMMENT	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
TITLE	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
JOURNAL	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
MEDLINE	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
FEATURES	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
source	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
Alignment Scores:	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
Pred. No.:	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
Score:	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
Percent Similarity:	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
Best local Similarity:	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
Query Match:	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
DB:	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
US-09-813-775C-18 (1-166) x BE663803 (1-444)	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
QY 28 G1GYSALAGLUNISQYSSerLeuAsnG1LethrValProASPThrLySValAsn 47	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
Db 152 GCGTGTGTCAGAAAGCGTCAGCTTCATATGAGATATCACTGTGCCAGACACCAAGTTAAC 211	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
QY 48 PheTyrAlaTrpLysAlaGAsn***Ser***GlnGlnAlaValGlnValATrGln 65	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
Db 212 TTCTATGCTGTCAGAAAGCGTCAGCTTCACCGCTTCCTCTTGCGAA 265	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
RESULT 4	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
CNS02F50	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
LOCUS	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
CNS02F50	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
LOCUS	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
CNS02F50	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
LOCUS	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
CNS02F50	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
LOCUS	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.	BE663803	BE663803.1	GI:10022630	Bos taurus	CCW.	EST.
CNS02F50	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence						

DEFINITION	Tetraodon nigroviridis genome survey sequence T7 end of clone 13H14 of library G from Tetraodon nigroviridis, genomic survey sequence.		
ACCESSION	AL194613		
VERSION	AL194613.1 GI:7832719		
KEYWORDS	GSS: genome survey sequence.		
SOURCE	Tetraodon nigroviridis.		
ORGANISM	Tetraodon nigroviridis		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.		
AUTHORS	1 (bases 1 to 968)		
TITLE	Roest-Crolius,H., Jalllon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fitzames,C., Wincker,P., Brothier,P., Quetier,F., Saurin,W. and Weissenbach,J.		
JOURNAL	Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 968)		
TITLE	Roest-Crolius,H., Jalllon,O., Dasilva,C., Fitzames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.		
JOURNAL	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis		
REFERENCE	Unpublished		
AUTHORS	3 (bases 1 to 968)		
TITLE	Genoscope.		
JOURNAL	Direct Submission		
COMMENT	Submitted (12-APR-2000)		
FEATURES	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon .		
SOURCE	Location/Qualifiers		
	1..968		
	/organism="Tetraodon nigroviridis"		
	/db_xref="taxon:99883"		
	/clone="13H14"		
	/clone_1b="G"		
	/note="Genoscope sequence ID : COAG131DD07LP1-end : T7"		
BASE COUNT	243 a 231 c 242 g 244 t 8 others		
ORIGIN			
Alignment Scores:			
Pred. NO.:	0.000191	Length:	968
Score:	120.00	Matches:	46
Percent Similarity:	40.12%	Conservative:	23
Best Local Similarity:	26.74%	Mismatches:	41
Query Match:	14.23%	Indels:	63
DB:	17	Gaps:	5
US-09-813-775C-18 (1-166) x CNS02F50 (1-968)			
QY 1	AlAProPArLeuLlECyASpSerArYValLeuGluArGYrLeuLeuGluAlaLys 20		
Db 304	TCCGCCCTGAGACCCATCTGTGACCTGAGAGTCCTCAACACATTTCATCAAGAGAGCCCAA 363		
QY 21	GLuAlaGlu-----AsnLe-ThrThrGly----- 28		
Db 364	GAGCGCGAAGCCGCCCATGTGTAGCGTTAACGTTGGAAAAAAGCAGMAAGAAAAAGAGCT 423		
QY 29	-----Cy 29		
Db 424	GTGACGCTAACAGAGTTTGTCCCGCTAACGTCGACGACCTCTACTCTCTTTAGAAAGACTG 483		
QY 29	salagluHisCysSerLeuEngluAsnLeThrValProAspThrLysValaAsnPhety 49		
Db 484	TAGAGAGGATGCAACCCCTGTCTGAGAGTCTGTGTGTCCGCCAGAACAGATGATTTTGA 543		
QY 49	rAlaTrP---LysArGAsn***----- 55		
Db 544	CGTCGTGGAGAAAGAAAAAGCT--AAGAMGMAAAGCTGTGTGTGTATATTCCTGCTGTC 602		

OY 56 -----Ser***GinginalaValGluValr 64
 Db 603 TTGGGATTCACCTCTTACTCTCTTGGCAAGCATCAGGAGGCGGAGAAAGTGA 662
 OY 64 polinglaleualseuSerglualaleuarglglnalaleuValasne 84
 Db 663 GTCTGGCTTGTGCTCCTACACGAGGCTTCACGACTTCCTCACCACCGC 722
 OY 84 rserglnprpGluProleuglnleuHisValasplysAlaValsergilyleuargse 104
 Db 723 ACC-----CTGCACAGCCACATAGCAACGCCGCTCCGCAACCTGCTCAG 767
 OY 104 rleuthrthrleuAargAlaleuGlyAlaGln 115
 Db 768 CCGTCAAGCGTGTGCTGGCAGCTCAACATCCAG 801
 RESULT 5
 LOCUS BM279790 365 bp mRNA linear EST 30-DEC-2001
 DEFINITION zahn505 Zebrafish Adult hearts cDNA library Danio rerio cDNA clone
 zahn505 5', mRNA sequence.
 ACCESSION BM279790
 VERSION BM279790.1 GI:17999056
 KEYWORDS EST.
 SOURCE zebrafish.
 ORGANISM Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
 ; Cyprinidae; Danio.
 REFERENCE 1 (bases 1 to 365)
 AUTHORS Ton, C., Dempsey, A.A., Hwang, D.M. and Liew, C.C.
 TITLE Identification and Characterization of Expressed Sequence Tags from
 Zebrafish Adult Hearts cDNA Library
 JOURNAL Unpublished (2001)
 COMMENT Contact: Liew CC
 Brigham and Women's Hospital
 Harvard Medical School
 75 Francis St. Boston, MA 02115, USA
 Tel: 6177328915
 Fax: 6179750995
 Email: cliw@rics.bwh.harvard.edu
 PCR Primers
 FORWARD: 5' GCCAAGCTCGAATTACCCCTCAGTAAAGG 3'
 BACKWARD: 5' CCACTGAATGTATACGACTCAGTATAGGCG 3'
 Insert Length: 365 Std Error: 0.00
 Seq primer: T3.
 FEATURES
 source
 1. 365
 Location/Qualifiers
 /organism="Danio rerio"
 /db_xref="taxon:7955"
 /clone="zahn505"
 /clone_lib="Zebrafish Adult hearts cDNA library"
 /dev_stage="adult"
 /note="Organ: heart; Vector: Lambda ZAP Express; Site_1:
 EcoRI; Site_2: XhoI; mRNA was purified from zebrafish
 adult hearts. cDNA was synthesized using a XhoI-Oligo dT
 adaptor-primer. EcoRI adaptors were ligated, followed by
 digestion with XhoI, for directional cloning into
 pre-digested lambda ZAP Express vector."
 BASE COUNT 69 a 89 c 95 g 112 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0.00262 length: 365
 Score: 106.00 Matches: 21
 Percent Similarity: 65.96% Conservative: 10
 Best Local Similarity: 44.68% Mismatches: 16
 Query Match: 12.57% Indels: 0
 DB: 13 Gaps: 0
 US-09-813-775c-18 (1-166) x BM279790 (1-365)

OY 1 AlAProPtoArgleuIlleCysasSperArgValleuGluArgTyrleuGluAlaLys 20
 Db 222 TCCCATATACGCCCATCTGACCTGCCTGCCTGCACATTTATTAAGAGGACATG 281
 OY 21 GluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerleuAsnGluAsnIleThr 40
 Db 282 GATGACAGAGGCTGCTATGAGAACTTGTAAGACGATTCGACATTGCAACGACGACT 341
 OY 41 ValProAspThrLysValAsn 47
 Db 342 GTTCCTTTGACGAGAGTCAT 362
 RESULT 6
 LOCUS BM315886
 DEFINITION f666c09.y1 sugano SUD adult male Danio rerio cDNA clone 5612560 5'
 similar to SW:ERO_MOUSE P07321 EXTHROPOLEITIN PRECURSOR. [1] ;,
 mRNA sequence.
 ACCESSION BM315886
 VERSION BM315886.1 GI:18050231
 KEYWORDS EST.
 SOURCE zebrafish.
 ORGANISM Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
 ; Cyprinidae; Danio.
 REFERENCE 1 (bases 1 to 611)
 AUTHORS Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy
 , S., Hillier, L., Kucaba, T., Martin, D., Beck, C., Wylie, T., Underwood
 , K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person, B.,
 Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E.,
 Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.
 and Wilson, R.
 TITLE WashU Zebrafish EST Project 1998
 JOURNAL Unpublished (1998)
 COMMENT Other-ESTs: f666c09.x1
 Contact: Stephen L. Johnson
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: zbrfish@wustl.edu
 Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA
 Sequencing by: Washington University Genome Sequencing Center Clone
 distribution information can be found through the I.M.A.G.E.
 Consortium/BLNI, send email to: info@image.llnl.gov
 Seq primer: T3 ET from Amersham
 High quality sequence stop: 526.
 FEATURES
 source
 1. 611
 Location/Qualifiers
 /organism="Danio rerio"
 /db_xref="taxon:7955"
 /clone="5612560"
 /clone_lib="Sugano SUD adult male"
 /sex="male"
 /tissue_type="whole body"
 /dev_stage="adult"
 /lab_host="DH10B (phage resistant)"
 /note="Vector: pME18S-FL3; Site_1: DraIII (CACCATGTG);
 Site_2: DraIII (CACTGTG); 1st strand cDNA was primed
 with an oligo(dT) primer (ATGCGCCCTTTTCTTTTCTTTTCTTTT);
 double-stranded cDNA was ligated to a DraIII adaptor
 (TGTGGCTCTAGTG), digested and cloned into distinct DraIII
 sites of the pME18S-FL3 vector (5' site CACGTGTG, 3' site
 CACCATGTG). XhoI should be used to isolate the cDNA
 insert. Size selection was performed to exclude fragments
 <1.5kb. Library constructed and donated by Dr. Sumio
 Sugano (University of Tokyo Institute of Medical Science).
 Custom primers for sequencing: 5' end primer
 CTCTGCTTAAAGCTGCG and 3' end primer
 CGACCTGCAGCTGAGCACA."
 BASE COUNT 156 a 172 c 144 g 139 t
 ORIGIN

	/ldc_host="DH10B (Life Technologies) (T1 phage resistant) /note=Organ: eye: Vector: pRT73-Pac (Pharmacia) with a modified polylinker. Site_1: EcoR I; Site_2: Not I; UT-E-Ed1 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pRT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are: fetal eyes, AGAATCAGA ; lens, CGATTAGCGA; eye anterior segment, AATCCGCAT; optic nerve, CCAATTAGTG; retina, CCGGG; Retina foreal and Macular, GTCC, RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."	
BASE COUNT	135 a 176 c 170 g 111 t	1 others
ORIGIN		
Alignment Scores:		
Pred. No.:	0.323	Length: 593
Score:	92.00	Matches: 21
Percent Similarity:	95.45%	Conservative: 0
Best Local Similarity:	95.45%	Mismatches: 1
Query Match:	10.91%	Indels: 1
DB:	14	Gaps: 0
US-09-813-775C-18 (1-166) x BM93J1504 (1-593)		
QY	145 TyiserAsnpheLeuAg9GLysleudLysleuTYrThrglyGUAlAcysArqThrGly 164	
Dd	3 TACTGCCAATTTCCHCCGGGGAAGTGG-AAACGTGTACACAGGGAGGCGCTGCAGCACAGG 61	
QY	165 AspArg 166	
Dd	62 GACACA 67	
RESULT 8	657 bp mRNA linear EST 30-APR-2002	
B0185758		
LOCUS	UT-E-Ed1-a)p-f.-03-0-UI.r1 UT-E-Ed1 Homo sapiens cDNA clone	
DEFINITION	UT-E-Ed1-a)p-f.-03-0-UI 5', mRNA sequence.	
ACCESSION	B0185758	
VERSION	B0185758.1 GI:20361309	
KEYWORDS	EST.	
SOURCE	human. Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 657) Bonaldo,M.F., Lennon,G. and Soares,M.B. Normalization and subtraction: two approaches to facilitate gene discovery Genome Res. 6 (9), 791-806 (1996)	
AUTHORS		
TITLE		
JOURNAL	genome Res. 6 (9), 791-806 (1996)	
MEDLINE	97044477	
COMMENT	Contact: Soares, MB Program for Rat Gene Discovery and Mapping University of Iowa 451 Eckstein Medical Research Building Iowa City, IA 52242, USA Tel.: 319 335 8250 Fax: 319 335 9565 Email: msoares@blue.weeg.uiowa.edu Tissue Procurement: Dr. Gregg Hagaman CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com). Seq primer: M13 REVERSE. Location/Qualifiers I..657	
FEATURES		
source	I..657	

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-E-EJ1-ajp-f-03-0-UI"
/clone_lib="UI-E-EJ1"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPE and
choroid"
/dev_stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site.1: EcoR I; Site.2: Not I;
UI-E-EJ1 is a subtracted cDNA library constructed
according to Honaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dt)18 tail. The
sequence tags for this library are: fetal eyes, AGATCAAGA
; lens, CGATTAGCGA; eye anterior segment, AATGCCGAT;
optic nerve, CCATTAGTG; retina, CCGCG; Retina Foveal and
Macular, CTCG; RPE and Choroid, ACTCA. This library was
created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI)."

BASE COUNT 166 a 189 c 176 g 123 t 3 others
ORIGIN

Alignment Scores:

Pred. No.: 0.376 Length: 657
Score: 92.00 Matches: 21
Percent Similarity: 95.45% Conservative: 0
Best Local Similarity: 95.45% Mismatches: 1
Query Match: 10.91% Indels: 1
DB: 14 Gaps: 0

US-09-813-775c-18 (1-166) x BQ185758 (1-657)

QY 145 TTTSeranpheleuarglylsleuylsleutyrrhgluaglaacysarrghrgly 164
LOCUS |||||||
DEFINITION 3 TACTTCATTTCCTCGGGGAAAGTG-AGCTGTACACAGGGAGGCTCGACGACAGG 61
DB |||||||

QY 165 Asparag 166
DB 62 GACAGA 67

RESULT 9

AZ413715 598 bp DNA linear GSS 03-OCT-2000
LOCUS AZ413715
DEFINITION 1M0197C23R Mouse 10kb plasmid UDCG1M library Mus musculus genomic
clone UDCG1M0197C23 R, DNA sequence.
AZ413715
ACCESSION AZ413715.1 GI:10537728
VERSION GSS.
KEYWORDS house mouse.
SOURCE Mus musculus
ORIGIN Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 598)

REFERENCE

AUTHORS

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

TITLE

JOURNAL

COMMENT

Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT
84112 USA
Tel: 801 585 5606

Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0197 row: C column: 23
Seq primer: CACACAGAAACACCTATGACAC
Class: plasmid ends
High quality sequence stop: 598.
Location/Qualifiers

FEATURES

source

1..598

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UDCG1M0197C23"
/clone_lib="Mouse 10kb plasmid UDCG1M library"
/sex="male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pMD42rv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g1147321149b/AP129072.1) a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT 151 a 129 c 144 g 174 t
ORIGIN

Alignment Scores:

Pred. No.: 1.21 Length: 598
Score: 87.50 Matches: 39
Percent Similarity: 37.76% Conservative: 15
Best Local Similarity: 27.27% Mismatches: 44
Query Match: 10.38% Indels: 45
DB: 17 Gaps: 7

US-09-813-775c-18 (1-166) x AZ413715 (1-598)

QY 9 SerArgValLeuGluArgTyrLeuGluAlaLysGluAlaGlnAlaThrThrgly 28
LOCUS |||
DEFINITION 185 AGCAATTCGCTTGACAACTTACTTTGTATTTTCGAAATTTGACATCGARCA 244
DB |||

QY 29 CysAlaGlu---HisCysSerLeuAsnGlnAlaThrAlaProAspThrLysAlaSn 47
DB 245 GCTCTCAGAGGACATTGC-----ATTGCG 268

QY 48 PheTyrAlaTyr-----LysArgAsn**Ser**GlnGlnAla 60
DB 269 TTTCTTCAGTGGGATGCTCATTCAGTTCGAAACCTAGTTGCAATATAGTGGGTCT 328

QY 61 ValGluValTTPGln-----GlyLeuAlaLeuLeuSerGlu 72
DB 329 GTCCAGCTTGTGCTTGTGAAAGTGGCCCGACATAGTGTGATACGTCCTCCCAAT 388

QY 73 AlaValLeuArgGlyGlnAlaLeuValAsnSer-----SerGlnProTyr----- 88
DB 389 TCAAAGTATTAATGGTGAAGCTCTTTGTACGACACATCGACACTGCTGTC 448

QY 89 -----GluProLeu-----GlnLeuHisVal 95
DB 449 TGTGAGACTGACAGCGGCTGAAAGCCTAATGATTACCAAGAGAAATGCTGACACA 508

QY 96 AspLysAlaValSerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGln 115

```

Db      509 GTTACAGCCTCAGCCGGGTGGTGGCTCAGACCTCTAATCCAGCCCTTGGAGGACAG 568
Qy      116 LysGluAla 118
        |||||
Db      569 AAGGAGCA 577

RESULT 10
A2981594      721 bp  DNA  linear  GSS 27-APR-2001
LOCUS      2M0262N06F Mouse 10kb plasmid UUGC2M library Mus musculus genomic
DEFINITION  clone UUGC2M0262N06 F, DNA sequence.
ACCESSION  A2981594
VERSION    A2981594
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus.
REFERENCE  Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 721)
            Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhasern,A.
            and Wright,D., Weiss,R.
            Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
            Unpublished (2000)
            Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert length: 10000 Std error: 0.00
            Plates: 0262 row: N column: 06
            Seq primer: CGTGTAAACGACGCGCGT
            Class: plasmid ends
            High quality sequence stop: 721.
            Location/Qualifiers
                1..721
                /organism="Mus musculus"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /clone="UUGC2M0262N06"
                /clone_lib="Mouse 10kb plasmid UUGC2M library"
                /sex="Female"
                /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
                /note="Vector: PMD42nv; Purified genomic DNA from M.
                musculus C57BL/6J (female) was obtained from the Jackson
                Laboratory Mouse DNA Resource
                (http://www.jax.org/resources/documents/dnares/). The DNA
                was hydrodynamically sheared by repeated passage through a
                0.005 inch orifice at constant velocity. The sheared DNA
                was blunt end-repaired with T4 DNA polymerase and T4
                polynucleotide kinase. Adaptor oligonucleotides were
                ligated to the blunt ends in high molar excess. The
                adaptor DNA was purified and size-selected for a 9.5 to
                10.5 kb range using preparative agarose gel
                electrophoresis. Vector DNA was prepared from a derivative
                of PMD42 (g1147321149b/AF129072.1), a copy-number
                inducible derivative of plasmid R1. The vector was ligated
                with adaptors complementary to the insert adaptors and
                purified. The sheared, adaptor mouse DNA was annealed to
                adaptor vector DNA, and transformed into
                chemically-competent E. coli XL10-Gold (Stratagene) cells
                and selected for ampicillin resistance."

BASE COUNT      193 a      172 c      160 g      196 t
ORIGIN
Alignment Scores:      2.15      Length:      721
Pred. No.:

```

```

Score:      86.50      Matches:      31
Percent Similarity:      45.74%      Conservative:      12
Best Local Similarity:      32.98%      Mismatches:      41
Query Match:      10.26%      Indels:      10
DB:      17      Gaps:      3

US-09-813-775c-18 (1-166) x A2981594 (1-721)
Qy      48 PhetYrAlATrPlySaRgAsn**Ser**GlnGlnAlaValGluValTrpGlnGlyLeu 67
        |||      |||      |||      |||      |||      |||      |||
Db      721 TTACGCCGATGGAGATGTGAGATGCGTCATGCCAAGAGACTCTAATTTGGCACTTTCTG 662
Qy      68 Ala-----LeuLeuSerGluAlaValLeuArgGlyGlnAlaLeuValAsn 83
        ::      |||      ::      |||      ::      |||      ::
Db      661 TCAGCAATTCCTTCAGATGACATACACAGGTTACACACTGAACCTCTGTAGTAGAA 602
Qy      84 SerSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSerGlyLeuArg 103
        |||||      |||||      |||
Db      601 CTCAGCCAGGCTGGGAACTTCATCTCTC-----CTCTGTGGGGTAACTTA 554
Qy      104 SerLeuThrThrLeuArgAlaLeu-GlyAlaGlnLysGluAlaLeuSerProProAs 123
        |||      |||      |||      |||      |||      |||      |||
Db      553 GAAACATGCACGCCCTTGGAATGCCATACACCTCTGAGAGAGAAAGA---AGCCCCCTGGC 497
Qy      123 pAlaAlaSerAlaAlaProLeuArgThrIleThrAlaAsp 136
        |||      |||      |||      |||      |||      |||      |||
Db      496 TACCTGTTCTGGAAGCGCTCTCAGAGACATCAGTGGAGAC 457
RESULT 11
BM072805      759 bp  mRNA  linear  EST 13-NOV-2001
LOCUS      MEST53-H09.T3 ISDN4-TN Zea mays cDNA clone MEST53-H09 3', mRNA
DEFINITION  sequence.
ACCESSION  BM072805
VERSION    BM072805.1 GI:16916346
KEYWORDS   EST.
SOURCE     Zea mays.
ORGANISM   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
            1 (bases 1 to 759)
            Qiu,F., Cui,F., Guo,L., Ashlock,D.A., Wen,T.J. and Schnable,P.S.
            Expressed Sequence Tags from B73 Maize Seedlings and Silks
            Unpublished (2001)
            Contact: Patrick S. Schnable
            Schnable Laboratory
            Iowa State University
            6405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA
            Tel: 515-294-0975
            Fax: 515-294-2299
            Email: schnable@iastate.edu
            Individual basecall and confidence value were assigned using the
            Phred software,
            (<http://depts.washington.edu/ventures/collabtr/direct/index.htm#b
            rt>). Overall sequence quality assessment and vector trimming were
            conducted using the Lucy software (<http://www.tigr.org/softlab/>).
            Lucy parameters were set to ensure an overall trimmed quality of
            97.5% or better without any vector fragments in the chosen
            high-quality region of each sequence. Low-quality bases between the
            poly-T and the high-quality region were replaced with N's to serve
            as spacers.
            PCR primers
            FORWARD: Forward PCR primer sequence, primer T7-1 (AA TAC GAC TCA
            CTA TAG)
            BACKWARD: Backward PCR primer sequence, primer T3 (ATT AAC CCT CAC
            TAA AG)
            Seq primer: primer T3 (ATT AAC CCT CAC TAA AG).
            Location/Qualifiers
                1..759
                /organism="Zea mays"
                /cultivar="B73"
                /db_xref="taxon:4577"
FEATURES
source

```


Category	Count	Percentage
a	182	17.2%
c	172	16.3%
t	123	11.7%
others	2	0.2%

Alignment Scores:

Pred. No.: 4.11 Length: 619
 Score: 83.50 Matches: 20
 Percent Similarity: 95.45% Conservative: 1
 Best Local Similarity: 90.91% Mismatches: 1
 Query Match: 9.91% Indels: 1
 DB: 14 Gaps: 0

US-09-813-775C-18 (1-166) x BM931641 (1-619)

OY 145 TyrSerAspPheLeuArgGlyLysLeuLysLeuTyrThrGlyAlaLacysArgThrGly 164
 |||||
 Db 3 TACTCCAAATTCCTCCGGGGGAAAGT--GAAGTGTACACAGGGAGGCGCTGACGAGACAGG 60

OY 165 AsparG 166

Db 61 GACAGA 66

RESULT 14

AZ572463 656 bp DNA linear GSS 15-MAY-2001
 LOCUS 306PvC10 Pv MBN #30 Plasmodium vivax genomic 3', DNA sequence.
 DEFINITION AZ572463
 ACCESSION AZ572463.1 GI:13985579

KEYWORDS GSS.
 SOURCE malaria parasite P. vivax.
 ORGANISM Plasmodium vivax

REFERENCE 1 (bases 1 to 656)
 AUTHORS Carlton, J.M.-R. and Dame, J.B.
 TITLE The Plasmodium vivax and P. berghei gene sequence tag projects
 JOURNAL Parasitol. Today (Regul. Ed.) 16 (10), 409 (2000)
 COMMENT Contact: Dame JB
 Dept. of Pathobiology, College of Veterinary Medicine
 University of Florida
 2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
 Tel: 352 392 4700
 Fax: 352 392 9704
 Email: damej@mail.ufl.edu
 Seq primer: M13(-20) forward
 Class: Shotgun.

FEATURES

Source

Location/Qualifiers
 1..656
 /organism="Plasmodium vivax"
 /strain="Salvador I (Collins, W. 1972, J. Parasitol. 69,
 497-598)"
 /db_xref="taxon:5855"
 /clone_id="Pv MBN #30"
 /dev_stage="asexual blood forms"
 /lab_host="Saimiri boliviensis"
 /note="Vector: pBluescript SK(+) vector DNA, phagemid
 excised from lambda ZAP; Site 1: EcoR V; Site 2: EcoR V;
 Host leukocytes were extracted from P. vivax infected
 blood using the following methods: first, infected blood
 was activated by the addition of 0.5 ml of ADP (40mg/ml)
 per 10 ml blood. Then blood was passed over a column of
 acid washed 0.1 mm glass beads, then through a Plasmidpur
 filter, followed by passage through a column of pre-wet
 Whatman CFI powder (1:2 ratio volume of blood to CFI),
 and finally centrifuged through a 50% Percoll density
 cushion. Purified DNA was digested with mung bean nuclease
 in the presence of 44% formamide at 500C as described
 (Vernick, K.D., Imberski, R.B., and McCutchan, T.F. 1988.
 Nucleic Acids Research 16:6883-6896). Digested DNA was
 blunt-ended using T4 DNA polymerase and size fractionated
 over a Sepharose CL-2B column. Fractions in the size range
 500bp-4kb were ligated into the Eco RV site of pBluescript
 SK(+), and E. coli XL-10 Gold transformed with the
 ligation mixture."
 BASE COUNT 101 a 200 c 195 g 157 t 3 others
 ORIGIN

Alignment Scores:

Pred. No.: 6 Length: 656
 Score: 82.50 Matches: 37
 Percent Similarity: 39.72% Conservative: 19
 Best Local Similarity: 26.24% Mismatches: 48
 Query Match: 9.79% Indels: 37
 DB: 17 Gaps: 6

US-09-813-775C-18 (1-166) x AZ572463 (1-656)

OY 59 GlnAlaValGluValProGlnGlyLeuAlaLeu----- 69

Db 98 CGCGTCATGAGGTGTGGGTGGTGTGCTGTGCTAGGGGGTGGATGATCGGCTACCGGC 157

OY 70 LeuSerGluAlaValLeuArgGlyGlnAlaLeuLeuValAsnSerSerGlnProTrpGlu 89

Db 158 CTCCTCACCGCTATGCTACCGCCCATGCTGTGCTGCATTGCGACCATGTGGCTTCAAG 217

OY 90 Pro---LeuGlnLeuHisValAspLys-----AlaValSerGlyLeuArg 103

Db 218 CCGCGCGTCCAGGGGCAATACCCCTTTCGCGGTTTAACTTGCCGCGCGCTCCGCTG 277

OY 104 SerLeuThrThrLeuLeuArgAlaLeu-----GlyAlaGlnLysGlnAlaLeuSer 120

Db 278 GCCCTTTCGCGCGCCCTTCCAAAGCTTCGATGATACCCCTGCTCCAGAGAAAGAGG 337

OY 121 ProProAspAlaAlaSerAlaAlaProLeuArgThrIleThr-----AlaAspThrPhe 138

Db 338 CCCCCCTTTCGAGGAGCAATTCACCTGATGATGCGGCTGCATGGCGAGATCGCTT 397

OY 139 ArgGlyLeuPheArgVal----- 144

Db 398 CCACAGTTCGTGCTGTGTGCTCCACTTCTTTCAGAAACGAGTACATACGTGTGAGG 457

OY 145 -----TyrSerAspPheLeuArgGlyLysLeuLysLeuTyrThrGlyAlaLacys 161

Db 458 AATTAAGCTTACGCGACCTTAAAGGGGAAATATCATTTATACAGGGGAGGAGATTG 517

OY 162 Arg 162

Db 518 GCG 520

RESULT 15

CNS02175 1035 bp DNA linear GSS 12-MAY-2000
 LOCUS CNS02175
 DEFINITION Tetraodon nigroviridis genome survey sequence PUC-ori end of clone
 224N04 of library G from Tetraodon nigroviridis, genomic survey
 sequence.

ACCESSION AL176522.1 GI:7814579
 VERSION AL176522.1
 KEYWORDS GSS, genome survey sequence.
 SOURCE Tetraodon nigroviridis.
 ORGANISM Tetraodon nigroviridis

REFERENCE 1 (bases 1 to 1035)
 AUTHORS Roest-Crollius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C.,
 Bernot, A., Fizames, C., Winkler, P., Brotlier, P., Queller, F.,
 Saurin, W., and Weissenbach, J.
 TITLE Human gene number estimate provided by genome wide analysis using
 Tetraodon nigroviridis DNA sequence

JOURNAL Unpublished
 2 (bases 1 to 1035)
 AUTHORS Roest-Crollius, H., Jallion, O., Dasilva, C., Fizames, C., Fisher, C.,
 Bouneau, L., Billault, A., Queller, F., Saurin, W., Bernot, A., and
 Weissenbach, J.

JOURNAL Unpublished
 3 (bases 1 to 1035)
 REFERENCE Genoscope.

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 8, 2003, 03:43:33 ; Search time 180.334 Seconds
(without alignments)
2072.993 Million cell updates/sec

Title: US-09-813-775c-18

Perfect score: 843
Sequence: 1 APRRLICDSRVLELYLEAK.....NFLRGLKLYTGECAGRTGDR 166

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+pn.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US09813775/runat.07012003.153120.23784/app.query.fasta_1.718
-DB=N.Geneseq.101002 -OPMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -DOOPT=0
-LOOPTXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09813775@cgn.1.1.0@runat.07012003.153120.23784 -NCPU=6 -ICPU=3
-NO_XDPXY -NO_MAP -LARGEOUTER -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMECUT=120
-NARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6 -FCGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
N.Geneseq.101002.*
1: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	825	97.9	498	21 ABL50878	Modified erythropo
2	825	97.9	501	19 AAV31031	Human erythropoiet
3	825	97.9	514	22 AAD06893	Human erythropoiet
4	825	97.9	533	6 AAN50346	Human recombinant
5	825	97.9	533	19 AAV30963	ECPEPO gene used in
6	825	97.9	551	6 AAN50345	Human recombinant
7	825	97.9	551	19 AAV30970	SCPEPO gene based o
8	825	97.9	582	21 AAA46697	DNA encoding a hum
9	825	97.9	582	22 AAH46972	Human erythropoiet
10	825	97.9	585	22 AAC66882	Chimpanzee erythro
11	825	97.9	614	24 ABL59290	Nucleotide sequenc
12	825	97.9	616	20 AAX77361	Polynucleotide seq
13	825	97.9	616	21 AAQ90393	DN encoding synthe
14	825	97.9	625	15 AAQ74760	Human prepro-eryth
15	825	97.9	629	24 ABL59289	Nucleotide sequenc
16	825	97.9	629	24 ABL59291	Nucleotide sequenc
17	825	97.9	724	9 AAN81554	EPO 140B encoding
18	825	97.9	788	17 AAT31529	Human erythropoiet
19	825	97.9	788	17 AAT31529	Human erythropoiet
20	825	97.9	823	18 AAT64847	Human erythropoiet
21	825	97.9	994	13 AAQ24281	IL-3:Epo short, re
22	825	97.9	1015	13 AAQ24281	Epo:IL-3 short, re
23	825	97.9	1051	13 AAQ24284	IL-3:Epo Flex, rec
24	825	97.9	1072	13 AAQ24285	Epo:IL-3 Flex, rec
25	825	97.9	1158	20 AAX25701	Human erythropoiet
26	825	97.9	1245	8 AAN70360	Sequence encoding
27	825	97.9	1255	7 AAN60519	CDNA of clone lamb
28	825	97.9	1342	7 AAN60513	CDNA of clone lamb
29	825	97.9	1789	16 AAQ92296	Erythropoietin cDN
30	825	97.9	3100	24 AAN86161	Nucleotide sequenc
31	825	97.9	10524	20 AAX77355	Polynucleotide seq
32	825	97.9	10524	21 AAA90387	Plasmid psinRep5 E
33	822	97.5	1239	18 AAT93979	DNA for oligopepti
34	821	97.4	468	21 AAA07253	Human Epo cDNA seq
35	820	97.3	498	21 AAA48373	Non-glycosylated e
36	820	97.3	501	21 ABL50863	Modified erythropo
37	820	97.3	570	21 ABL50864	Modified erythropo
38	820	97.3	582	21 ABL50865	Modified erythropo
39	817	96.9	975	19 AAV70367	Human flt3 ligand
40	817	96.9	1020	19 AAV70367	Human flt3 ligand
41	816	96.8	498	21 ABL50862	Modified erythropo
42	812	96.3	500	21 AAA29493	Erythropoietin cod
43	812	96.3	500	21 AAA29494	Erythropoietin cod
44	812	96.3	521	11 AAQ06473	Human erythropoiet
45	811	96.2	579	22 AAT71883	Erythropoietin gen

ALIGNMENTS

RESULT 1	ABL50878	ABL50878 standard; DNA; 498 BP.
ID	ABL50878;	
AC	ABL50878;	
XX		
XX		
DT	21-JUN-2002 (first entry)	
XX		
DE	Modified erythropoietin related gene sequence #5.	
XX		
KW	Modified erythropoietin; EPO; gene; ds.	
XX		
OS	Unidentified.	
FH	Key	Location/Qualifiers
FT	CDS	1..498
FT		/tag a
FT		/partial

```

FT      /product= "Modified erythropoietin protein"
FT      /note= "no start or stop codons given"
XX      KR145802-B1.
XX      PD      01-AUG-1998.
XX      PF      31-MAY-1994; 94KR-0012082.
XX      PR      31-MAY-1994; 94KR-0012082.
XX      PA      (GIDS ) LG CHEM CO LTD.
XX      PI      Kim C, Song Y, Lee T;
XX      DR      WPI: 2000-234250/20.
XX      P-PSDB: ABB07030.
XX      PT      MODIFIED ERYTHROPOIETIN GENE AND EXPRESSION VECTORS THEREOF -
XX      PS      Disclosure; Page 14; 15pp; Korean.
XX      CC      The present invention describes modified erythropoietin (EPO) genes
XX      CC      and expression vectors comprising the genes. The present sequence
XX      CC      represents a gene sequence from the present invention.
XX      SQ      Sequence 498 BP; 133 A; 103 C; 115 G; 147 T; 0 other;

Alignment Scores:
Pred. No.:      8.6e-84      Length:      498
Score:          825.00      Matches:      162
Percent Similarity: 97.59%      Conservative: 0
Best Local Similarity: 97.59%      Mismatches: 4
Query Match:    97.86%      Indels:      0
DB:             21      Gaps:         0

US-09-813-775c-18 (1-166) x ABL50878 (1-498)

QY      1 AAlaProAArgLeuIleCysAspSerArgValLeuGluArgTyrLeuLeuGluAlaLys 20
DB      1 GCTCCACCAAGATGATCTGTGATTCAGAGCTTGGAAGATCTTATTTGGAAGCTAAG 60
QY      21 GluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
DB      61 GAAGGTGAAGCATCACCACAGTGGTGTCTGTAACACTGTCCTGACGAAACAAATCACT 120
QY      41 ValProAArgThrLysValAsnPhetYrAlaTrpLysArgAsn***Ser**GlnGlnAla 60
DB      121 GTCCAGATACCAAGTAACTTCAACCTTGGAAGACAAAGAGAGAGTGGTCAACAGCT 180
QY      61 ValGluValITrpgInglyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
DB      181 GTCGAAAGTGGCAAGGTTTGGCTTGTGCTATCTGAAGCTGCTTCAAGAGGTCACACTTA 240
QY      81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSer 100
DB      241 TTGGTTAACTCTTTCACACCTTGGGAAACATTGCAATGACAGTGAAGAGCTGTCTCT 300
QY      101 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGluAlaIleSer 120
DB      301 GGTCTAAATCTCTTAACCTCTGTTGAGAGCTTTGGGTGCTCAAAAGGAACTATCTCT 360
QY      121 ProProAAspAlaIleSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLys 140
DB      361 CCTCCAGATGCTGTTCGCGCTCCATTTGAAACTATCAGTCTGATATCTTTCGAAGAG 420
QY      141 LeuPheArgValTyrSerAsnPhetLeuArgGlyLysLeuLysLeuTyrThrGlyGluAla 160
DB      421 TTGTTCACAGTCTTAACCTCTTTCGAGAGTAGTGAAGTGTGTACACCGGTGAAGCC 480
QY      161 CysArgThrGlyAspArg 166
DB      481 TGTAGAACTGGTGAATACA 498

```

```

RESULT 2
AAV31031
ID      AAV31031 standard; DNA; 501 BP.
XX      AC      AAV31031;
XX      DT      12-OCT-1998 (first entry)
XX      DE      Human erythropoietin template DNA.
XX      KW      Erythropoietin receptor agonist; EPO; human; anaemia;
XX      KW      haematopoietic deficiency; red blood cell; erythroid progenitor;
XX      KW      bone marrow suppression; ss.
XX      OS      Homo sapiens.
XX      PN      WO9818926-A1.
XX      PD      07-MAY-1998.
XX      PF      23-OCT-1997; 97WO-US18703.
XX      PR      25-OCT-1996; 96US-0034044.
XX      PA      (SEAR ) SEARLE & CO G D.
XX      PI      Feng Y, McWhorter CA, Summers N;
XX      DR      WPI: 1998-272221/24.
XX      P-PSDB: AAW58404.
XX      PT      Human erythropoietin receptor agonist polypeptide - used to
XX      PT      stimulate the production of red blood cells in a patient.
XX      PS      Example 1; Page 88-89; 112pp; English.
XX      CC      This DNA sequence, encoding a human erythropoietin (EPO)
XX      CC      polypeptide (see AAW58404), can be used as a template for the
XX      CC      production of claimed nucleic acid molecules (see AAV30971-V31030)
XX      CC      encoding novel circularly permuted EPO receptor agonists. PCR
XX      CC      amplifications or tandem duplication methods can be used to
XX      CC      rearrange the DNA sequence, and the novel agonists are expressed
XX      CC      in transformed or transfected host cells. Claimed EPO receptor
XX      CC      agonists (see AAW58413-72) have a new N-terminus and C-terminus,
XX      CC      with the original C-terminal polypeptide joined to the N-terminal
XX      CC      polypeptide directly or via a linker. They are used in claimed
XX      CC      methods for stimulating the production of haematopoietic cells, for
XX      CC      selective ex vivo expansion of erythroid progenitors, and for
XX      CC      treating patients having haematopoietic disorders.
XX      SQ      Sequence 501 BP; 108 A; 152 C; 145 G; 96 T; 0 other;

Alignment Scores:
Pred. No.:      8.67e-84      Length:      501
Score:          825.00      Matches:      162
Percent Similarity: 97.59%      Conservative: 0
Best Local Similarity: 97.59%      Mismatches: 4
Query Match:    97.86%      Indels:      0
DB:             19      Gaps:         0

US-09-813-775c-18 (1-166) x AAV31031 (1-501)

QY      1 AAlaProAArgLeuIleCysAspSerArgValLeuGluArgTyrLeuLeuGluAlaLys 20
DB      1 GCCCACCACCCCTCATCTGTGACAGCCGAGTCTGGAGAGGTAACCTTGTGAGGCCAAG 60
QY      21 GluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
DB      61 GAGGCCGAGATATCAGACGAGGCTGTCTGACACCTGCACTGAAATGAAATATATCACT 120
QY      41 ValProAArgThrLysValAsnPhetYrAlaTrpLysArgAsn***Ser**GlnGlnAla 60

```



```

Db 121 GTCCAGACACCAAGTAAATTTCTATGCTTGGAGAGAGTGGAGTCCGGCAGCAGGCC 180
Oy 61 valGluValITrPcInGlyLeuAlaLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
Db 181 GTAGAAAGTCTGGAGAGGCGCTGGCTGTGTGGAAAGCTGTCTGGGGGCGAGGCCCG 240
Oy 81 LeuValAsnSerSerGlnProITrpgIuProLeuGlnLeuHisValAspLysAlaValSer 100
Db 241 TTGGTCACACTTCCCGAGCCGTGGAGCCCTGCACCTCATGTGATAAAGCCGTCAGT 300
Oy 101 GtyleuArSerIeuThrIeuArGAlaLeuGtlyAlaGlnLysGluAlaIleSer 120
Db 301 GGGCTTGGAGGCTCACCTCTGCTTGGGCTTGGAGGCCAGAGAGGCAATCTCC 360
Oy 121 ProProAspAlaAlaSerAlaIaProLeuArgThrIleThraIAspThrPheArgLys 140
Db 361 CCTCCAGATGGCGCCCTCAGCTGCTCCACTCCGACAAATACATCTCAGACTTTCCGCAA 420
Oy 141 LeuPheArgValTyrSerAsnPheLeuArgGlyLysLeuLysLeuTyrThrGlyGluAla 160
Db 421 CTCCTCCGAGCTACTACCAATTTCCCTCGGGAAGCTGAACTGTAAACACAGCGAGCCG 480
Oy 161 CysArgThrGlyAspArg 166
Db 481 TCGAGGACAGGGGACAGA 498

RESULT 3
AAD06893
ID AAD06893 standard: cDNA; 514 BP.
XX
AC AAD06893:
XX
DT 06-AUG-2001 (first entry)
XX
DE Human erythropoietin (EPO) cDNA.
XX
KW Human; erythropoietin; EPO; antianaemic; nephrotrophic; anti-HIV;
KW vaccine; haemostatic; Immunoglobulin; Ig; EPO deficient disease;
KW anaemia; renal failure; Human Immunodeficiency Virus; HIV;
KW haematopoietic growth factor; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 8..508
FT FT /*tag= a
FT FT /product= "Human erythropoietin"
FT FT /note= "CDS does not include start codon"
FT FT /partial
FT FT 7
FT FT misc_feature
FT FT /*tag= b
FT FT /note= "Original base is changed to "r" which
FT FT results in increased expression"
FT FT 40
FT FT misc_feature
FT FT /*tag= c
FT FT /note= "Original base is changed to "g" which
FT FT results in increased expression"
FT FT 88
FT FT misc_feature
FT FT /*tag= d
FT FT /note= "Original base is changed to "c" which
FT FT results in increased expression"
FT FT 121
FT FT misc_feature
FT FT /*tag= e
FT FT /note= "Original base is changed to "c" which
FT FT results in increased expression"
FT FT 127
FT FT misc_feature
FT FT /*tag= f
FT FT /note= "Original base is changed to "c" which
FT FT results in increased expression"
FT FT 130
FT FT misc_feature
FT FT /*tag= g
FT FT /note= "Original base is changed to "g" which
FT FT results in increased expression"

```

```

FT FT misc_feature 133
FT FT /*tag= h
FT FT /note= "Original base is changed to "r" which
FT FT results in increased expression"
FT FT 145
FT FT misc_feature
FT FT /*tag= i
FT FT /note= "Original base is changed to "g" which
FT FT results in increased expression"
FT FT 175
FT FT misc_feature
FT FT /*tag= j
FT FT /note= "Original base is changed to "r" which
FT FT results in increased expression"
FT FT 178
FT FT misc_feature
FT FT /*tag= k
FT FT /note= "Original base is changed to "c" which
FT FT results in increased expression"
FT FT 196
FT FT misc_feature
FT FT /*tag= l
FT FT /note= "Original base is changed to "g" which
FT FT results in increased expression"
FT FT 283
FT FT misc_feature
FT FT /*tag= m
FT FT /note= "Original base is changed to "a" which
FT FT results in increased expression"
FT FT 304
FT FT misc_feature
FT FT /*tag= n
FT FT /note= "Original base is changed to "g" which
FT FT results in increased expression"
FT FT 344
FT FT misc_feature
FT FT /*tag= o
FT FT /note= "Original base is changed to "g" which
FT FT results in increased expression"
FT FT 400
FT FT misc_feature
FT FT /*tag= p
FT FT /note= "Original base is changed to "c" which
FT FT results in increased expression"
FT FT 489..490
FT FT misc_feature
FT FT /*tag= q
FT FT /note= "Original bases are changed to "g" and "c" which
FT FT results in increased expression"
FT FT 509..514
FT FT misc_feature
FT FT /*tag= r
FT FT /note= "Original base are changed to "CTCGAG" which
FT FT results in increased expression"

WO200136489-A2.
XX
PN 03-NOV-2000; 2000WO-EP10843.
XX
PD 12-NOV-1999; 99US-0164855.
XX
PD 25-MAY-2001.
XX
PF 03-NOV-2000; 2000WO-EP10843.
XX
PR 12-NOV-1999; 99US-0164855.
XX
PA (MERCK ) MERCK PATENT GMBH.
XX
PI Hartmann A, Brandt S, Rieke E, Sobel C, Lo K, Way JC, Gillies S;
XX WPI: 2001-367563/38.
XX DR P-PSDB; AA02641.
XX
XX Novel modified erythropoietin forms such as fusion proteins, comprising
XX FC portion of an immunoglobulin molecule and a target molecule having
XX PT the biological activity of erythropoietin forms
XX PS
XX Example 1; Page 22; 58pp; English.
XX
XX The present sequence is a cDNA encoding human erythropoietin (EPO)
XX CC mature protein. The coding sequence for mature EPO contains modified
XX CC codons to optimise translation. EPO has improved biological activity and
XX CC an extended serum half life greater than 20 hours. The present invention
XX CC relates to modified EPO forms such as fusion proteins comprising a FC
XX CC portion of an immunoglobulin (Ig) molecule and an EPO molecule (Fc-EPO).
XX CC The FC portion is fused covalently through its C-terminus directly or

```

CC indirectly to the EPO molecule, and where the Fc portion as well as EPO
 CC portion may be modified or mutated. The invention also relates to non-
 CC fused EPO molecules which have a pattern of cysteines or disulphide
 CC bonding which is distinct from human or animal EPO. Pharmaceutical
 CC compositions containing EPO are useful in the treatment of EPO deficient
 CC diseases such as anaemia, renal failure, HIV infection, blood loss and
 CC chronic disease that can be treated with haematopoietic growth factor.

XX Sequence 514 BP; 106 A; 159 C; 152 G; 97 T; 0 other;

Alignment Scores:

Pred. No.:	8.96e-84	Length:	514
Score:	825.00	Matches:	162
Percent Similarity:	97.59%	Conservative:	0
Best Local Similarity:	97.59%	Mismatches:	4
Query Match:	97.86%	Indels:	0
DB:	22	Gaps:	0

US-09-813-775c-18 (1-166) x AAD06893 (1-514)

```

QY 1 AlaProAaGLeuIleCysAspSerArgValLeuGluArgTyrLeuGluAlaLys 20
DB 8 GCCCACCACGCGCTCATCTGTGACAGCGAGTGTGAGAGGTACTCTTGAGGCCAAG 67
QY 21 GluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
DB 68 GAGCGCGAATATCACCACCGCGCTGTCTGACACTGCTGACCTTGAAGACATCACC 127
QY 41 ValProAspThrLysValAsnPhetYrAlaTrpLysArgAsn**Ser**GlnGlnAla 60
DB 128 GTGCTGACACCAAGTGAATTTCTATGCTGTGAGAGAGATGAGAGTGGCCACAGGCC 187
QY 61 ValGluValITrpgInglyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
DB 188 GTAGAGATGTGGACAGGCGCTGGCTGTGTCGAAAGCTGTCTCGGGGCCAGCCCTG 247
QY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSer 100
DB 248 TTGTCACACTCTCCACACCGCGTGGAGGCCCTGCAACTGCACTGATGAATAAACCCGTGAGT 307
QY 101 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGlnAlaLeu 120
DB 308 GGCCTTCACAGCTCACCACATCTGCTTGGGCTCTGGAGGCCACAGAGGCCATCTCC 367
QY 121 ProProAspAlaIleSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLys 140
DB 368 CCTCCAGATGGGCGCTGCTGCTCCCTCCGACAAATCAGTGTGACACTTTCGCAAA 427
QY 141 LeuPheArgValITyrSerAsnPheLeuArgGlyLysLeuLysLeuTyrThrGlyGlnAla 160
DB 428 CTCTTCGAGTCTACTCAATTTCTCTCGGGGAAAGCTGAAGCTGTACACAGGGAGGCC 487
QY 161 CysArgThrGlyAspArg 166
DB 488 TGCCGGACAGGGGACAGA 505

```

RESULT 4
 AANS0346
 ID AANS0346 standard; DNA: 533 BP.

XX AC AANS0346;
 XX 01-JAN-1980 (first entry)
 DE Human recombinant erythropoietin expressed in *Escherichia coli*.
 XX Erythropoietin; red blood cell; erythrocyte; anaemia; blood;
 KW disorder; ds; *Escherichia coli*.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH mat_peptide 27..528
 FT

FT /*tag- a

XX W08502610-A.

XX 20-JUN-1985.

XX 11-DEC-1984; 84WO-US02021.

XX 30-NOV-1984; 84US-0675298.

XX 13-DEC-1983; 83US-0561024.

XX 21-FEB-1984; 84US-0582185.

XX 28-SEP-1984; 84US-0655841.

XX (KIRI-) KIRIN-AMGEN INC.

XX WPI: 1985-159229/26.

XX P-PSDB: AAP50299.

XX New polypeptide having properties of erythropoietin - is prepd.
 by cultivation of transformed eucaryotic or procaryotic host

XX Disclosure; Page 72; 113pp: English.

XX Human erythropoietin encoded by this sequence is essential for red
 blood cell formation and is used for the diagnosis and treatment of
 blood disorders such as anaemia. Large amounts of EPO may be obtained
 using recombinant DNA techniques in contrast to small amounts
 obtained from plasma and urine. This sequence is expressed in E
 coli. See also AANS0345, AANS0347-50 and AAP50298, AAP50300-PS0301.

XX Sequence 533 BP; 131 A; 134 C; 138 G; 130 T; 0 other;

Alignment Scores:

Pred. No.:	9.4e-84	Length:	533
Score:	825.00	Matches:	162
Percent Similarity:	97.59%	Conservative:	0
Best Local Similarity:	97.59%	Mismatches:	4
Query Match:	97.86%	Indels:	0
DB:	6	Gaps:	0

US-09-813-775c-18 (1-166) x AANS0346 (1-533)

```

QY 1 AlaProAaGLeuIleCysAspSerArgValLeuGluArgTyrLeuGluAlaLys 20
DB 30 GCTCCGCGCGCTGTGATCTGCGACCTGAGAGTGTGAACTTACTCTGGAAGCTTAA 89
QY 21 GluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
DB 90 GAGCTGAAAACATCACCACCTGCTGCTGCTGACACTGTCTTGAACGAAACATTTACG 149
QY 41 ValProAspThrLysValAsnPhetYrAlaTrpLysArgAsn**Ser**GlnGlnAla 60
DB 150 GTACCAGACACCAAGTGAATCTTACGCTTGGAAGCTATGAAGTGTGCACCAACCA 209
QY 61 ValGluValITrpgInglyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
DB 210 GTTGAAGTTTGGACAGGCTGTGCACTGTGAGCGGCGTGTACTGCGGCGAGGCACTG 269
QY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSer 100
DB 270 CTGTGTAACCTCTCTCAGCCCTGGGACCCGCTGCAGCTGTCTGTGACAAAGCATATCT 329
QY 101 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGlnAlaLeu 120
DB 330 GGCCTGAGATCTGTGACATCTGCTGTGCTGTGCTGTGCTGTGACAGAGGATATCTCT 389
QY 121 ProProAspAlaIleSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLys 140
DB 390 CCGCGGATGTGCTGATGTGACCCGCTGCTGACATCATCTGATGATACCTTCGCAAA 449
QY 141 LeuPheArgValITyrSerAsnPheLeuArgGlyLysLeuLysLeuTyrThrGlyGlnAla 160
DB 450 CTGTTTCTGTATACTCTAACTCTCTGCTGTGATGAAGTGAAGCTGTATATGCGGAACCA 509

```

```

OY 161 CysArgThrGlyAspArg 166
DB 510 TGCCTACTGTGACCGC 527

RESULT 5
AAV30963
ID AAV30963 standard; DNA; 533 BP.
XX
AC AAV30963;
XX
DT 11-SEP-1998 (first entry)
XX
DE ECEPO gene used in the construction for human EPO.
XX
KW Human; erythropoietin; EPO; bone marrow; reticulocyte; red blood cell;
KM expression; CHO; chinese hamster ovary cell; diagnosis; blood disorder;
KK ds.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN AU688723-B.
XX
PD 19-FEB-1998.
XX
PF 02-DEC-1997; 97AU-0046867.
XX
PR 02-DEC-1997; 97AU-0046867.
XX
PA (KIRI ) KIRIN AMGEN INC.
XX
PI Lin F;
XX
DR WPI: 1998-261957/24.
XX
PT Recombinant human erythropoietin - potentially useful for diagnosis
PT and treatment of blood disorders
XX
PS Example 11; Page 68; 100pp; English.
XX
CC The present sequence represents the ECEPO gene which resembles the
CC amino terminus of human erythropoietin (EPO). The present invention
CC describes recombinant human EPO which causes bone marrow cells to
CC increase production of reticulocytes or red blood cells, where the
CC polypeptide is the product of expression in CHO (Chinese hamster ovary)
CC cells of an exogenous DNA sequence encoding human EPO. EPO is
CC potentially useful in the diagnosis and treatment of blood disorders
CC characterised by low or defective red blood cell production.
XX
SQ Sequence 533 BP; 131 A; 134 C; 138 G; 130 T; 0 other;

Alignment Scores:
Pred. No.: 9.4e-84 Length: 533
Score: 825.00 Matches: 162
Percent Similarity: 97.59% Conservative: 0
Best Local Similarity: 97.59% Mismatches: 4
Query Match: 97.86% Indels: 0
DB: 19 Gaps: 0

US-09-813-775C-18 (1-166) x AAV30963 (1-533)
OY 1 AlaProPArgLeuIleCysAspSerArgValLeuGluArgTyrLeuLeuGluAlaLys 20
DB 30 GCTCCGCGCGTCTGTGATGCGACTCGAGAGTCTGGAACGTTACTGCTGGAAGCTAA 89
OY 21 GluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
DB 90 GAAGCGAAACATCAACCACTGCTGTGCTGAACACTGTTCTTGAACGAAACATTACG 149
OY 41 ValProAspThrIleValAsnPhetYrAlaTrpIleAsn***Ser***GlnGlnAla 60
DB 150 GTACCAACACCAAGGTTAACTTCTACGCTTGGAACGTAATGCAAGTTGGTCAACAGCA 209

```

```

OY 61 ValGluValTrpGlnGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
DB 210 GTTGAAGTTTGGCAGGGGTCTGGCAGCTGCTGAGAGGAGGCTGACTGGCTGGCAGGACAG 269
OY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSer 100
DB 270 CTGGTAAACCTCTCTCAGCGCGTGGAAACCGCTTCACCTGCTGATGTGACAAACGACATCTC 329
OY 101 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlnValAlaGlnIleGlnAlaIleSer 120
DB 330 GGCTTGAGATCTCTGACTACTCTGCTGCGTCTGGGTCCACAGAAAGGAGCTATCTCT 389
OY 121 ProProAspAlaAlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLys 140
DB 390 CCGCCGGATCTCTCATCTGCTGTGCACCGCTGCTGACATCATCTCTATACCTTCGCCGAA 449
OY 141 LeuPheArgValIlyrSerAsnPhetLeuArgGlyLysLeuLysLeuTyrThrGlyGluAla 160
DB 450 CTGTTTCGTGTATACCTTCTGCGGTAAACTGAACTGATATCTGCGGAACGA 509
OY 161 CysArgThrGlyAspArg 166
DB 510 TGCCTACTGTGACCGC 527

RESULT 6
AAN50345
ID AAN50345 standard; DNA; 551 BP.
XX
AC AAN50345;
XX
DT 01-JAN-1980 (first entry)
XX
DE Human recombinant erythropoietin expressed in Saccharomyces
DE cerevisiae.
XX
KW Erythropoietin; red blood cell; erythrocyte; anaemia; blood;
KM disorder; ds; Saccharomyces cerevisiae.
XX
OS Homo sapiens.
XX
FH Key location/Qualifiers
FT mat_peptide complement (13..547)
FT /tag= a
XX
PN W08502610-A.
XX
PD 20-JUN-1985.
XX
PF 11-DEC-1984; 84MO-US02021.
XX
PR 30-NOV-1984; 84US-0675298.
PR 13-DEC-1983; 83US-0561024.
PR 21-FEB-1984; 84US-0582185.
PR 28-SEP-1984; 84US-0655841.
XX
PA (KIRI-) KIRIN-AMGEN INC.
XX
DR WPI: 1985-159229/26.
DR P-PSDB; AAP50298.
XX
PT New polypeptide having properties of erythropoietin - is prepd.
PT by cultivation of transformed eucaryotic or procaryotic host
XX
PS Disclosure; Page 82; 113pp; English.
XX
CC Human erythropoietin encoded by this sequence is essential for red
CC blood cell formation and is used for the diagnosis and treatment of
CC blood disorders such as anaemia. Large amounts of EPO may be obtained
CC using recombinant DNA techniques in contrast to small amounts
CC obtained from plasma and urine. This sequence is expressed in S.
CC cerevisiae. See also AAN50346-50 and AAP50299-P50301.
XX

```

SO Sequence 551 BP; 157 A; 112 C; 129 G; 153 T; 0 other;

Alignment Scores:

Pred. No.:	9,82e-84	Length:	551
Score:	825.00	Matches:	162
Percent Similarity:	97.59%	Conservative:	0
Best Local Similarity:	97.59%	Mismatches:	4
Query Match:	97.86%	Indels:	0
DB:	6	Gaps:	0

US-09-813-775c-18 (1-166) x AAN50345 (1-551)

```

OY 1 AAlaProAlaGLeuILECysAspSerArgValLeuGluArgTyrLeuGluAlaLys 20
    |||||
DB 16 GCTCCACCAAGATTGATCTGTGACTCGAGAGCTTTGGAAAGATCTTTGGAACTAAA 75
OY 21 GluAlaGluAsnILEthrThrGlyCysAlaGluHisCysSerLeuAsnGluSnILEthr 40
    |||||
DB 76 GAAAGCTGAAACATCAACCACTGCTGTGCTGACACCTGCTTTGAAACGAAACATTACG 135
OY 41 ValProAspThrLysValAsnPhetYrAlaTrpLysArgAsn***Ser***GlnGlnAla 60
    |||||
DB 136 GTACCAAGACACCAAGTTACTTCTACGCTTGGAAACGATATGGAAGTGTGCAACAGCT 195
OY 61 ValGluValITrPGlnGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
    |||||
DB 196 GTTGAAGTTTGGCAAGTTTGGCTTGTATCTGAAGCTGTTTGAAGCTCAACGCTTGG 255
OY 81 LeuValAsnSerSerGlnProITrPGluProLeuGlnLeuHisValAspLysAlaValSer 100
    |||||
DB 256 TTGGTTAACTCTTCTCAACCATGGAACCATTTGCAATTTGACGTGATTAAGCCGCTCT 315
OY 101 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGluAlaIleSer 120
    |||||
DB 316 GGTTTGAGATCTTTGACACTTTGTTGAGAGCTTTGGGCTCAAAAGAACCACTTCC 375
OY 121 ProProAspAlaIleSerAlaIleAlaProLeuArgThrIleThrAlaAspThrPheArgLys 140
    |||||
DB 376 CCACGAGACGCTGCTTGCCTCCATTTGAGAACCATTCACGTGATTAACCTTTCAGAAAG 435
OY 141 LeuPheArgValITyrSerAsnPhetLeuArgGlyLysLeuLysLeuTyrThrGlyGluAla 160
    |||||
DB 436 TTATTCAAGATTACTTCCAACTTCTTGAGAGGTAAATTGAAGTTGTACACGCGGTAGACC 495
OY 161 CysArgThrGlyAspArg 166
    |||||
DB 496 TGTAGAACTGGTGACAGA 513

```

RESULT 7

AAV30970 AAV30970 standard; DNA; 551 BP.

AC AAV30970;

DT 11-SEP-1998 (first entry)

DE SCFPO gene based on the human EPO gene with preferred yeast codons.

KW Human; erythropoietin; EPO; bone marrow; reticulocyte; red blood cell;

KW expression; CHO; chinese hamster ovary cell; diagnosis; blood disorder;

OS Synthetic.

OS Homo sapiens.

PN AU688723-B.

PD 19-FEB-1998.

PF 02-DEC-1997; 97AU-0046867.

PR 02-DEC-1997; 97AU-0046867.

PA (KIRI) KIRIN AMGEN INC.

PI Lin F;

DR WPI; 1998-261957/24.

PT Recombinant human erythropoietin - potentially useful for diagnosis

PS and treatment of blood disorders

PS Example 11; Page 78; 100pp; English.

The present sequence represents the SCFPO gene which was constructed as part of the assembly of human erythropoietin (EPO) with yeast preferred codons. The present invention describes recombinant human EPO which causes bone marrow cells to increase production of reticulocytes or red blood cells, where the polypeptide is the product of expression in CHO (chinese hamster ovary) cells of an exogenous DNA sequence encoding human EPO. EPO is potentially useful in the diagnosis and treatment of blood disorders characterised by low or defective red blood cell production.

SO Sequence 551 BP; 157 A; 112 C; 129 G; 153 T; 0 other;

Alignment Scores:

Pred. No.:	9,82e-84	Length:	551
Score:	825.00	Matches:	162
Percent Similarity:	97.59%	Conservative:	0
Best Local Similarity:	97.59%	Mismatches:	4
Query Match:	97.86%	Indels:	0
DB:	19	Gaps:	0

US-09-813-775c-18 (1-166) x AAV30970 (1-551)

```

OY 1 AAlaProAlaGLeuILECysAspSerArgValLeuGluArgTyrLeuGluAlaLys 20
    |||||
DB 16 GCTCCACCAAGATTGATCTGTGACTCGAGAGCTTTGGAAAGATCTTTGGAAGCTAAA 75
OY 21 GluAlaGluAsnILEthrThrGlyCysAlaGluHisCysSerLeuAsnGluSnILEthr 40
    |||||
DB 76 GAAAGCTGAAACATCAACCACTGCTGTGCTGACACCTGCTTTGAAACGAAACATTACG 135
OY 41 ValProAspThrLysValAsnPhetYrAlaTrpLysArgAsn***Ser***GlnGlnAla 60
    |||||
DB 136 GTACCAAGACACCAAGTTACTTCTACGCTTGGAAACGATATGGAAGTGTGCAACAGCT 195
OY 61 ValGluValITrPGlnGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
    |||||
DB 196 GTTGAAGTTTGGCAAGTTTGGCTTGTATCTGAAGCTGTTTGAAGCTCAACGCTTGG 255
OY 81 LeuValAsnSerSerGlnProITrPGluProLeuGlnLeuHisValAspLysAlaValSer 100
    |||||
DB 256 TTGGTTAACTCTTCTCAACCATGGAACCATTTGCAATTTGACGTGATTAAGCCGCTCT 315
OY 101 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGluAlaIleSer 120
    |||||
DB 316 GGTTTGAGATCTTTGACACTTTGTTGAGAGCTTTGGGCTCAAAAGAACCACTTTC 375
OY 121 ProProAspAlaIleSerAlaIleAlaProLeuArgThrIleThrAlaAspThrPheArgLys 140
    |||||
DB 376 CCACGAGACGCTGCTTGCCTCCATTTGAGAACCATTCACGTGATTAACCTTTCAGAAAG 435
OY 141 LeuPheArgValITyrSerAsnPhetLeuArgGlyLysLeuLysLeuTyrThrGlyGluAla 160
    |||||
DB 436 TTATTCAAGATTACTTCCAACTTCTTGAGAGGTAAATTGAAGTTGTACACGCGGTAGACC 495
OY 161 CysArgThrGlyAspArg 166
    |||||
DB 496 TGTAGAACTGGTGACAGA 513

```

RESULT 8

AAA46697 AAA46697 standard; DNA; 582 BP.

AC AAA46697;
 XX
 DT 25-SEP-2000 (first entry)
 XX
 DE DNA encoding a human erythropoietin polypeptide.
 XX
 KW Human; erythropoietin; EPO; inhibitor; nuclear factor-kappaB; NF-kappaB;
 XX multi-drug resistance gene; malignant hemopathy; solid tumour;
 XX malignant blood disease; leukaemia; lymphoma; solid cancer; ds.
 OS
 XX Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 1..582
 FT /*tag= a
 FT /product= "erythropoietin"
 XX
 PN MO200030587-A2.
 XX
 PD 02-JUN-2000.
 XX
 PF 24-NOV-1999: 99WO-FR02897.
 XX
 PR 25-NOV-1998: 98FR-0014858.
 XX
 PA (CNRS) CENT NAT RECH SCI.
 XX
 PI Hirsch F, Haeflner A;
 XX
 DR WPI: 2000-399901/34.
 XX P-PSDB; AA93638.
 DR
 XX
 PT Treatment of haematological or solid tumours using an inhibitor of the
 PT activation of nuclear factor-kappaB, particularly to prevent
 PT development of resistance to chemotherapeutics
 XX
 PS Claim 5; Page 28-29; 30pp; French.
 XX
 CC The present sequence encodes a human erythropoietin (EPO) polypeptide.
 CC The human growth hormone protein is used as an inhibitor of the
 CC activation of nuclear factor-kappaB (NF-kappaB). The inhibitor inhibits
 CC activation of NF-kappaB, and thus transcription of the multi-drug
 CC resistance gene (which contains binding sites for NF-kappaB within its
 CC regulatory regions). The inhibitors are used to produce pharmaceuticals
 CC which may be used in the treatment of malignant hemopathy or solid
 CC tumours. The inhibitors are especially used to treat malignant blood
 CC diseases (leukaemia, lymphoma) and solid cancers (of breast or ovary).
 XX
 SQ Sequence 582 BP; 113 A; 181 C; 170 G; 118 T; 0 other;
 Alignment Scores:
 Pred. No.: 1,06e-83 Length: 582
 Score: 825.00 Matches: 162
 Percent Similarity: 97.59% Conservative: 0
 Best Local Similarity: 97.59% Mismatches: 4
 Query Match: 97.86% Indels: 0
 DB: 21 Gaps: 0
 US-09-813-775C-18 (1-166) x AAA46697 (1-582)
 QY 1 A1AProPARGleu1ecysAspSerArgValLeuCluArgTyrLeuLgualAlaLys 20
 DB 82 GCGCCACAGCGCTCATGTGTGACACCGAGCTCTGGAGAGTCTTGGAGGCCAAG 141
 QY 21 GluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
 DB 142 GAGCGCGAATAATATCATCGACGAGGCTGTCTGCAACTGCGAGCTTGAATGCAATATCACT 201
 QY 41 ValProAspThrLysValAsnPhetyrAlaIrrLysArgAsn**Ser**GlnGlnAla 60
 DB 202 GTCCCGACAGACCAAAAGTAAATTTCTATGCTGTGAGAGAGTGAAGTGGCGGACGAGGCC 261
 QY 61 ValGluValTrpGlnIleuAlaLeuLeuSerGlnAlaValLeuArgGlyGlnAlaLeu 80

DB 262 GTAGAAAGTCTGGCAGGGCCCTGGCCCTGCTCGGAAGCTGCTCGGGGCGCAGGCCCTG 321
 QY 81 LeuValAsnSerSerGlnProTrrpGluProLeuGlnLeuHisValAspLysAlaValSer 100
 DB 322 TTGGTCACTCTTCCAGCCGCTGGAGACCCCTGACGTGATGTGATTAAGCCGTCAGT 381
 QY 101 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGlnAlaIleSer 120
 DB 382 GGCCTTGCAGCCTCAGACCACTGCTGCGGCTCTGGAGCCAGGAAGAGCATCTCC 441
 QY 121 ProProAspAlaAlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLys 140
 DB 442 CCTTCAGATGCGGCTCAGCTGCTCCAGCAATCACTGCTGACACTTTCGCGAAA 501
 QY 141 LeuPheArgValTyrSerAsnPhelLeuArgGlyLysLeuLysLeuTrrThrGlyGlnAla 160
 DB 502 CTCTCCGAGTCTACTTCATTTTCTCTCCGGGGAAGCTGAAGCTGTACACAGGGAGGCC 561
 QY 161 CysArgThrGlyAspArg 166
 DB 562 TGCAGACAGGAGGACAGA 579
 RESULT 9
 AAH46972
 ID AAH46972 standard; cDNA; 582 BP.
 XX
 AC AAH46972;
 XX
 DT 29-OCT-2001 (first entry)
 XX
 DE Human erythropoietin (EPO) cDNA sequence.
 XX
 KW Transgenic; pig; human; erythropoietin; EPO; milk; PMSG; hCG;
 KW chorionic gonadotrophic hormone; WAP promoter; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..582
 FT /*tag= a
 FT /product= "EPO"
 FT sig_peptide 1..81
 FT /*tag= b
 FT mat_peptide 82..579
 FT /*tag= c
 XX
 PN WO200159074-A1.
 XX
 PD 16-AUG-2001.
 XX
 PF 28-JUN-2000: 2000WO-KR00675.
 XX
 PR 14-FEB-2000: 2000KR-0006888.
 XX
 PA (KORE-) REPUBLIC KOREA.
 XX
 PI Chang W, Park J, Seong H, Min K, Yang B, Im G, Lee Y, Lee C;
 XX Kim J;
 DR WPI: 2001-514656/56.
 DR P-PSDB; AAB85573.
 XX
 PT Producing transgenic porcine that secretes human erythropoietin (hEPO)
 PT in milk by introducing vector comprising hEPO genome into fertilized
 PT eggs of porcine to which PMSG and hCG were administered, and developing
 PT progeny
 XX
 PS Claim 4; Fig 3; 21pp; English.
 XX
 CC The invention relates to producing transgenic pigs (P) that secrete
 CC human erythropoietin (hEPO) in milk. The method involves administering
 CC PMSG and human chorionic gonadotrophic hormone (hCG) into (P), collecting

CC fertilized eggs after mating, injecting expression vector containing a
 CC 2.6 kb WAP promoter, hEPO genome and SV40 poly A DNA into male pronuclei,
 CC transplanting them in surrogate mother pig and allowing it to give birth.
 CC The method provides transgenic porcine capable of secreting hEPO in their
 CC milk, thus producing the expensive useful medicine at a low cost with
 CC stability on a large scale, giving a contribution to the improvement of
 CC human health. The present sequence represents a base sequence for a human
 CC EPO cDNA incorporated into the genome of porcine.

XX Sequence 582 BP; 113 A; 181 C; 170 G; 118 T; 0 other;

Alignment Scores:

Pred. No.:	1.06e-83	Length:	582
Score:	825.00	Matches:	162
Percent Similarity:	97.59%	Conservative:	0
Best Local Similarity:	97.59%	Mismatches:	4
Query Match:	97.86%	Indels:	0
DB:	22	Gaps:	0

US-09-813-775c-18 (1-166) x AAC6682 (1-582)

```

OY 1 AAlAPrOArGLuEulIEcYsAsPSeArGVaLleuGluArGTyrLeuEugLuAlaLys 20
    |||
DB 82 GCGCCACACAGCGCTATCTGTGACAGCCGAGTCTCGAGAGGTACTCTTGAGGCCAAG 141
    |||
OY 21 GluAlaGluAsnIleThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
    |||
DB 142 GAGGCCGGAATATACAGACGCGGCTGTGCTGACACTGACGCTGATGATGAGAAATATCACT 201
    |||
OY 41 ValProAspThrIlyValAsnPhetyrAlaTrpIlyAsrAsn***Ser***GlnGlnAla 60
    |||
DB 202 GTCCCGAGACACCAAGTTAATTTCTATGCTGGAAGAGATGAGGTCGGCGCAGAGGCC 261
    |||
OY 61 ValGluValITrPGlnGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
    |||
DB 262 GTAGAAGTCTGGCAGAGCGCTGCTGCTGCTGGAAGCTGTCTGCGGCCAGGCCCTG 321
    |||
OY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspIlyAlaValSer 100
    |||
DB 322 TTGGTCAACTCTTCCACCGCGTGGAGGCCCGCGAGCTGATGATGAAGCCGTCAGT 381
    |||
OY 101 GlyLeuArgSerLeuThrThreLeuArgAlaLeuGlyAlaGlnIlyGlnAlaIleSer 120
    |||
DB 382 GCGCTTCGAGCTCAGCACCTGCTGCTGCTGCTGCGGAGCCAGAGAGGACCATCTCC 441
    |||
OY 121 ProProAspAlaAlaSerAlaAlaProLeuArgThrIleThraAlaAspThrPheArgLys 140
    |||
DB 442 CCTCCAGATGGCGCTCAGCTGCTCCACGACATCAGATCAGCTTCCGCAAA 501
    |||
OY 141 LeuPheArgValIlySerAsnPhetyrAlaTrpIlyGlyLysLeuIlyLeuTyrThrGlyGlnAla 160
    |||
DB 502 CTCCTCCGAGTCTACTCCAAATTTCTCGGGGAAAGCTGAAGCTGTACACAGGGAGGCC 561
    |||
OY 161 CysArgThrGlyAspArg 166
    |||
DB 562 TGCAGGACAGGGGACAGA 579
    |||
  
```

RESULT 10 AAC66882 ID AAC66882 standard; cDNA: 585 BP.

XX AAC66882;

XX 27-MAR-2001 (first entry)

DE Chimpanzee erythropoietin coding sequence SEQ ID NO: 3.

KW Chimpanzee; erythropoietin; EPO; hybridisation probe; gene therapy;

KW mapping; therapeutic agent; ss.

XX Pan sp.

XX WO200068376-A1.

XX 16-NOV-2000.
 XX
 XX 05-MAY-2000; 2000WO-US12370.
 XX
 XX 07-MAY-1999; 99US-0307307.
 XX 28-MAR-2000; 2000US-0307307.
 XX
 XX (GETH) GENENTECH INC.
 XX
 XX Desauvage F, Henner DJ;
 XX
 XX WPI: 2001-007393/01.
 XX
 XX Nucleic acids encoding chimpanzee erythropoietin, useful for treatment
 XX of e.g. anemia, also derived proteins, antibodies and modulators -
 XX Claim 2; Fig 2; 10pp; English.

CC The present invention provides the coding and protein sequences of
 CC chimpanzee erythropoietin (EPO). These sequences can be used in gene
 CC therapy, to block the activity of EPO, as hybridisation probes, in
 CC genetic and chromosome mapping and as therapeutic agents.

SQ Sequence 585 BP; 115 A; 185 C; 168 G; 117 T; 0 other;

Alignment Scores:

Pred. No.:	1.06e-83	Length:	585
Score:	825.00	Matches:	162
Percent Similarity:	97.59%	Conservative:	0
Best Local Similarity:	97.59%	Mismatches:	4
Query Match:	97.86%	Indels:	0
DB:	22	Gaps:	0

US-09-813-775c-18 (1-166) x AAC66882 (1-585)

```

OY 1 AAlAPrOArGLuEulIEcYsAsPSeArGVaLleuGluArGTyrLeuEugLuAlaLys 20
    |||
DB 82 GCGCCACACAGCGCTATCTGTGACAGCCGAGTCTCGAGAGGTACTCTTGAGGCCAAG 141
    |||
OY 21 GluAlaGluAsnIleThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
    |||
DB 142 GAGGCCGGAATATACAGACGCGGCTGTGCTGACACTGACGCTGATGATGAGAAATATCACT 201
    |||
OY 41 ValProAspThrIlyValAsnPhetyrAlaTrpIlyAsrAsn***Ser***GlnGlnAla 60
    |||
DB 202 GTCCCGAGACACCAAGTTAATTTCTATGCTGGAAGAGATGAGGTCAGGACGAGGCC 261
    |||
OY 61 ValGluValITrPGlnGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
    |||
DB 262 GTAGAAGTCTGGCAGAGCGCTGCTGCTGCTGCTGGAAGCTGTCTGCGGCCAGGCCCTG 321
    |||
OY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspIlyAlaValSer 100
    |||
DB 322 TTGGTCAACTCTTCCACCGCGTGGAGGCCCGCGAGCTGATGATGAAGCCGTCAGT 381
    |||
OY 101 GlyLeuArgSerLeuThrThreLeuArgAlaLeuGlyAlaGlnIlyGlnAlaIleSer 120
    |||
DB 382 GCGCTTCGAGCTCAGCACCTGCTGCTGCTGCTGCGGAGCCAGAGAGGACCATCTCC 441
    |||
OY 121 ProProAspAlaAlaSerAlaAlaProLeuArgThrIleThraAlaAspThrPheArgLys 140
    |||
DB 442 CCTCCAGATGGCGCTCAGCTGCTCCACGACATCAGATCAGCTTCCGCAAA 501
    |||
OY 141 LeuPheArgValIlySerAsnPhetyrAlaTrpIlyGlyLysLeuIlyLeuTyrThrGlyGlnAla 160
    |||
DB 502 CTCCTCCGAGTCTACTCCAAATTTCTCGGGGAAAGCTGAAGCTGTACACAGGGAGGCC 561
    |||
OY 161 CysArgThrGlyAspArg 166
    |||
DB 562 TGCAGGACAGGGGACAGA 579
    |||
  
```

RESULT 11

ABL59290
 ID ABL59290 standard; DNA; 614 BP.
 AC ABL59290;
 XX
 DT 07-OCT-2002 (first entry)
 XX
 DE Nucleotide sequence of a modified human erythropoietin (EPO).
 XX
 KW Human; erythropoietin; EPO; glycoprotein; reticulocyte production;
 KW red blood cell production; anaemia; chronic renal failure;
 KW acquired immunodeficiency syndrome; AIDS; cancer; bone marrow;
 KW committed erythroid progenitor; ss.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 14..604
 FT sig_peptide /*tag= a
 FT 14..94 /*tag= b
 FT /note= "secretion signal sequence"
 FT
 XX
 PM WO200249673-A2.
 XX
 PD 27-JUN-2002.
 XX
 PF 08-DEC-2001; 2001WO-EP1434.
 XX
 PR 20-DEC-2000; 2000EP-0127891.
 XX
 PA (HOFF) HOFFMANN IA ROCHE & CO AG F.
 XX
 PI Burg J, Engel A, Franze R, Hilger B, Schurig HE, Tischer W;
 PI Wozny M;
 PI
 XX
 DR WPI: 2002-566640/60.
 DR P-PSDB; ABB77902.
 XX
 PT Novel conjugate of erythropoietin glycoprotein with polyethylene
 PT glycol, useful for treating diseases correlated with anaemia in chronic
 PT renal failure patients and acquired immunodeficiency syndrome -
 XX
 PS Disclosure; Fig 4; 40pp; English.
 XX
 CC The present sequence encodes a modified human erythropoietin (EPO)
 CC protein. The EPO was extended at the N-terminal by a proteolytic
 CC cleavage site. It was used to produce conjugates of the invention. The
 CC specification describes a conjugate comprising an EPO glycoprotein having
 CC an N-terminal alpha-amino group, chosen from human EPO (hEPO) or its
 CC analogues (where hEPO is modified by addition of 1-6 glycosylation sites
 CC or a rearrangement of a glycosylation site). The glycoprotein is
 CC covalently linked to a poly(ethylene glycol) group. The EPO glycoprotein
 CC has in vivo biological activity of causing bone marrow cells to increase
 CC production of reticulocytes and red blood cells. The conjugate increased
 CC circulating half-life and plasma residence time, decreased clearance,
 CC increased clinical activity in vivo, improved potency and stability, when
 CC compared to unmodified EPO. The EPO conjugate is useful for preparing
 CC medicaments for the treatment and prophylaxis of diseases correlated with
 CC anaemia in chronic renal failure patients (CRF), acquired
 CC immunodeficiency syndrome (AIDS) and for treating cancer patients
 CC undergoing chemotherapy. It is also useful for treating patients by
 CC stimulating the division and differentiation of committed erythroid
 CC progenitors in the bone marrow.
 XX
 SQ Sequence 614 BP; 119 A; 198 C; 176 G; 121 T; 0 other;

Alignment Scores:
 Pred. No.: 1.13e-83 Length: 614
 Score: 825.00 Matches: 162
 Percent Similarity: 97.59% Conservative: 0
 Best Local Similarity: 97.59% Mismatches: 4

Query Match: 97.86% Indels: 0
 DB: 24 Gaps: 0
 US-09-813-775C-18 (1-166) x ABL59290 (1-614)
 QY 1 AlaProProArgLeuIleCysAspSerArgValLeuGlnArgTyrLeuLeuGlnAlaLys 20
 DB 104 GCCCCACACAGCCCTCATCTGTGTGACAGCCGAGTCTGTGAGAGGTACCTCTTGAGGCGCAAG 163
 QY 21 GluAlaGlnAsnIleThrThrGlyCysAlaGlnHisCysSerLeuAsnGlnAsnIleThr 40
 DB 164 GAGGCCGGAATATTCACGACGGGCTGTGTGAACACTGCACCTTGATATGATATACACT 223
 QY 41 ValProAspThrLysValAsnPhetYrAlaTyrLysArgAsn**Ser**GlnGlnAla 60
 DB 224 GTCCACAGACACCAAGTAAATTTCTATGCTGTGAGAGAGATGAGAGTCCGCGACAGCGCC 283
 QY 61 ValGluValTyrGlnGlyLeuAlaLeuLeuSerGlnAlaValLeuArgLysGlnAlaLeu 80
 DB 284 GTAGAAGTGTGGCAGGGCTGTGGCTGTGTGGAAGCTGTCTGTGGGGCCAGGGCCCTG 343
 QY 81 LeuValAsnSerSerGlnProTyrGlnProLeuGlnLeuHisValAspLysAlaValSer 100
 DB 344 TTGGTCAACTCTTCCACGCCGTGGAGCCCTGCACCTCATGTGATTAAGCCGTCACT 403
 QY 101 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGlnAlaIleSer 120
 DB 404 GGCCCTTCGACGCTCACCACTCTGCTTGGGGCTGTGGAGCCAGAGCAAGCAACATCTCC 463
 QY 121 ProProAspAlaAlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLys 140
 DB 464 CTCCTCAGATGGCGCTCACCTGTCTCCACATCATGCTGTACACTTTCGCCAAA 523
 QY 141 LeuPheArgValTyrSerAsnPhetLeuArgGlyLysLeuLysLeuThrGlyGlnAla 160
 DB 524 CTCTTCGAGTCTACTCCAAATTTCTCCGGGGAAGCTGAAGCTGTACACAGGGAGGCC 583
 QY 161 CysArgThrGlyAspArg 166
 DB 584 TGCAGGACAGGGGACAGA 601
 RESULT 12
 ID AAX77361
 AC AAX77361 standard; DNA; 616 BP.
 XX
 XX AAX77361;
 AC
 XX
 DT 09-AUG-1999 (first entry)
 XX
 DE Polynucleotide sequence of synthetic erythropoietin.
 XX
 KW Nucleic acid identification; exogenous protein; gene sorting;
 KW growth factor; membrane receptor; sindbis virus; erythropoietin; ss.
 XX
 OS Synthetic.
 OS
 XX
 PN WO9925876-A1.
 XX
 PD 27-MAY-1999.
 XX
 PF 17-NOV-1998; 98WO-US24520.
 XX
 PR 17-NOV-1997; 97US-0972218.
 XX
 PA (CYTO-) CYTOS BIOTECHNOLOGY GMBH.
 XX
 PI Bailey JE, Koller D, Orberger GH, Renner WA;
 PI
 XX
 DR WPI: 1999-357620/30.
 XX
 PT Isolating genes encoding proteins with selected properties, useful
 PT for identifying therapeutic agents or targets
 XX

PS Disclosure: Fig 15; 136pp; English.

XX
CC The invention relates to the identification of a recombinant nucleic acid
CC encoding an exogenous protein having a selected property. The method
CC comprises preparing a population of eukaryotic host cells, culturing the
CC cells under suitable conditions and identifying cells that contain the
CC recombinant nucleic acid. The method is used to sort genes according to
CC the type of proteins they express, and also to identify new ligand/
CC receptor interactions. Typical applications of the nucleic acid and the
CC exogenous protein are in isolation of new growth factors, cytokines,
CC membrane receptors, cytoplasmic, organelle or nuclear proteins, all of
CC which may be useful as therapeutic agents or therapeutic targets, e.g.
CC apoptosis promoting or tumour suppressing proteins, regulators of cell
CC proliferation or metabolic processes etc. The protein can also be used to
CC screen for specific modulators. The nucleic acid can also be used as
CC sources of therapeutic antisense or ribozyme sequences. The method allows
CC the protein (rather than a partial DNA sequence) to be isolated and,
CC since a wide range of cells can be used, they can be expressed with the
CC correct glycosylation pattern.

XX
SQ Sequence 616 BP; 118 A; 191 C; 184 G; 123 T; 0 other;

Alignment Scores:

Pred. No.:	1.14e-83	Length:	616
Score:	825.00	Matches:	162
Percent Similarity:	97.59%	Conservative:	0
Best Local Similarity:	97.59%	Mismatches:	4
Query Match:	97.86%	Indels:	0
DB:	20	Gaps:	0

US-09-813-775C-18 (1-166) x AAX77361 (1-616)

QY 1 AAlAPROAATGLeuIIeCysAspSerArgValLeuGluArgTyrLeuLeuGluAlaLys 20
DB 97 GCCCCACACGCGCTCATGTGACAGCCGAGCTCGGAGAGGTACTCTTGAGAGCCCAAG 156
QY 21 GluAlaGluAsnIleThrThrGlyCysAlaGluIuHsCysSerLeuAsnGluAsnIleThr 40
DB 157 GAGGCCGGAATATCATCAGCAGCGGCTGTGTGAACTGAGCTTGAATGAGAAATATCACT 216
QY 41 ValProAspThrLysValAsnPhenylAlaTrpLysArgAsn**Ser**GlnGlnAla 60
DB 217 GTCCACAGACACCAAGTTAATTCTATGCTGTGAGAGAGATGAGGTGGGCGACAGGCC 276
QY 61 ValGluValITrpgLgIleuAlaLeuSerGluAlaValLeuArgLysGlnAlaLeu 80
DB 277 GTAGAAAGTGTGGCAGGCGCTGGCTGTGTGGAAGCTGTCTCGGGGCGACGCGCTG 336
QY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHsValAspLysAlaValSer 100
DB 337 TTGGTCACACTCTCCACCGCGGGAGCCCTGCGAGCTGCATGTGATTAAGCCGCTCAAT 396
QY 101 GlyLeuAspSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGlnAlaLeuSer 120
DB 397 GGCCTTCGACGCTCAGACACTCTGCTGGGCTCTGGGAGCCCAAGAACCCATCTCC 456
QY 121 PropioAspAlaAlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLys 140
DB 457 CCTCAGATGGGGCTCAGCTGCTCCACTCCGACAAATCAGACGCTGACACTTCCGCAAA 516
QY 141 LeuPheArgValITyrSerAsnPhenylLeuArgLysLeuLysLeuThrThrGlyGlnAla 160
DB 517 CTCTTCGAGTCTACTCCAAATTTCTCCGGGGAAGCTGAAGCTGTACACAGGGAGGCC 576
QY 161 CysArgThrGlyAspArg 166
DB 577 TGCAAGACAGGGGACAGA 594

RESULT 13

AAA90393

ID AAA90393 standard; DNA; 616 BP.

XX
AC AAA90393;

XX
DT 10-JAN-2001 (first entry)

XX
DE DN encoding synthetic erythropoietin (EPO).

XX
KW Nucleic acid identification; exogenous protein; drug screening;

XX
KW recombinant expression; synthetic erythropoietin; EPO; ds.

XX
OS Synthetic.

XX
PN JP2000189173-A.

XX
PD 11-JUL-2000.

XX
PF 23-AUG-1999; 99JP-0236220.

XX
PR 17-NOV-1998; 98US-0193707.

XX
PR 17-NOV-1998; 98WO-US24520.

XX
PA (CYTO-) CYTOS BIOTECHNOLOGY GMBH.

XX
DR WPI; 2000-551637/51.

XX
PT Identifying a recombinant nucleic acid to identify and isolate various
XX cellular proteins, comprises culturing a composition comprising
XX eukaryotic host cells and identifying a cell comprising recombinant
XX nucleic acid

XX
PS Disclosure: Fig 15; 56pp; Japanese.

XX
CC The invention relates to the identification of a recombinant nucleic acid
CC encoding an exogenous protein having a selected property. The method
CC comprises preparing populations of eukaryotic host cells, where each cell
CC comprises an expression vector encoding a different exogenous protein.
CC The host cells are cultured under suitable conditions and the nucleic
CC acid which encodes the exogenous protein is identified. The method is
CC useful for the identification and isolation of proteins with a selected
CC property. Typical applications of the nucleic acid and the exogenous
CC protein are in isolation of new growth factors, cytokines, membrane
CC receptors, cytoplasmic, organelle or nuclear proteins, all of which may
CC be useful as therapeutic agents or therapeutic targets, e.g., pro-
CC apoptotic or tumour suppressing proteins, regulators of cell
CC proliferation or of metabolic processes. The protein can also be used to
CC screen for ligands and specific modulators of activity. The method of the
CC invention allows the direct cloning of full length cDNAs in one step. It
CC facilitates direct expression of the protein without the need to perform
CC further procedures such as subcloning and establishment of a cell line
CC for protein production. The method allows a protein of interest (rather
CC than a partial DNA sequence) to be isolated and, since a wide range of
CC cell types can be used, they can be expressed in a correctly folded and
CC glycosylated form. The present sequence represents DNA encoding a
CC synthetic erythropoietin (EPO) which was used in the method of the
CC invention. This patent is related to WO9925876.

XX
SQ Sequence 616 BP; 118 A; 191 C; 184 G; 123 T; 0 other;

Alignment Scores:

Pred. No.:	1.14e-83	Length:	616
Score:	825.00	Matches:	162
Percent Similarity:	97.59%	Conservative:	0
Best Local Similarity:	97.59%	Mismatches:	4
Query Match:	97.86%	Indels:	0
DB:	21	Gaps:	0

US-09-813-775C-18 (1-166) x AAA90393 (1-616)

QY 1 AAlAPROAATGLeuIIeCysAspSerArgValLeuGluArgTyrLeuLeuGluAlaLys 20
DB 97 GCCCCACACGCGCTCATGTGACAGCCGAGCTCGGAGAGGTACTCTTGAGAGCCCAAG 156
QY 21 GluAlaGluAsnIleThrThrGlyCysAlaGluIuHsCysSerLeuAsnGluAsnIleThr 40
DB 157 GAGGCCGGAATATCATCAGCAGCGGCTGTGTGAACTGAGCTTGAATGAGAAATATCACT 216


```

QY 41 ValProAspThrIysValAsnPhETyRAlaTrpIysArgAsn***Ser***GInGlnAla 60
DB 217 GTCCACAGACACCAAGTTAATTCTATGCTCGAAGAGCATGAGTGGGACAGAGGCC 276
QY 61 ValGluValTrpGlnGlyLeuAlaLeuLeuSerGlnAlaValLeuArgGlyGlnAlaLeu 80
DB 277 GTAGAAAGTGTGGCAGGGCCGTGGCCCTGTCTGGAAGCTGTCTGGGGCCAGAGCCCTG 336
QY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspIysAlaValSer 100
DB 337 TTGGTCAACTCTTCCCGCGGTGGAGCCCTGACGTGATGGATAAAGCCGTAGT 396
QY 101 GlyLeuArgSerLeuThrTrpLeuLeuArgAlaLeuGlyAlaGlnIysGlnAlaIleSer 120
DB 397 GGCCTTCGACGCTCACACACTGTCTGCGGCTGTGGAGCCCGAAGAAAGCCATCTCC 456
QY 121 ProProAspAlaAlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLys 140
DB 457 CCTCCAGATGCGGCGCTGAGTGTCCACCTCGAACAATCACTGCTGACACTTTCGGCAA 516
QY 141 LeuPheArgValTyrSerAsnPhelLeuArgGlyLysLeuLysLeuTyrThrGlyAla 160
DB 517 CTCTTCGAGTCTACTCAATTTCCTCCGGGAAAGCTGAAGCTGTACACAGGGAGGCC 576
QY 161 CysArgThrGlyAspArg 166
DB 577 TGCAGACAGGGGACAGA 594

RESULT 14
AAQ74760 standard; DNA: 625 BP.
XX ID AAQ74760 standard; DNA: 625 BP.
XX AC AAQ74760:
XX
XX 24-JUN-1995 (first entry)
XX
XX Human prepro-erythropoietin.
XX
XX Erythropoietin; therapeutic; ss.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX FT 1..21 /*tag= a
XX FT 22..102 /*tag= b
XX FT 103..600
XX CDS /*tag= c
XX
XX MO9425055-A.
XX
XX 10-NOV-1994.
XX
XX 29-APR-1994: 94MO-US04755.
XX
XX 29-APR-1993: 93US-0055076.
XX
XX (ABBO ) ABBOTT LAB.
XX
XX Devries PJ, Mellowitz BS, Meuth JL, Okasinski GF;
XX Schaefer VG;
XX
XX WPI: 1994-357906/44.
XX
XX P-PSDB: AAR65499.
XX
XX Erythropoietin analogues - useful for treatment of anaemia and
XX have enhanced erythropoietic effect.
XX
XX Disclosure: Page 38-39; 56pp; English.
XX
XX The synthetic DNA encoding human prepro-erythropoietin may be ligated

```

```

CC into an expression vector for erythropoietin expression in a CHO
CC cell culture. Site-directed mutagenesis may be used in the
CC construction of Epo analogues with improved activity, which may be
CC used in pharmaceutical compositions for inducing erythropoiesis and
CC treating anaemia.
XX
XX SQ Sequence 625 BP; 123 A; 195 C; 179 G; 128 T; 0 other;

Alignment Scores:
Pred. No.: 1,16e-83 Length: 625
Score: 825.00 Matches: 162
Percent Similarity: 97.59% Conservative: 0
Best Local Similarity: 97.59% Mismatches: 4
Query Match: 97.86% Indels: 0
DB: 15 Gaps: 0

US-09-813-775c-18 (1-166) x AAQ74760 (1-625)
QY 1 AlaProAspThrIysValAsnPhETyRAlaTrpIysArgAsn***Ser***GInGlnAla 20
DB 103 GCCCACCACGCGCTCATATGTAAGTCCGAGTCCGAGAGTACTCTTGGAGGCCAAG 162
QY 21 GluAlaGlnAsnIleThrTrpGlyCysAlaGlnHisCysSerLeuAsnGlnIleThr 40
DB 163 GAGCCGAGAAATATTAGAGGGGCTGTGCTGAGCAGCTGGAATGAGAAATATCACT 222
QY 41 ValProAspThrIysValAsnPhETyRAlaTrpIysArgAsn***Ser***GInGlnAla 60
DB 223 GTCCACAGACACCAAGTTAATTCTATGCTCGAAGAGCATGAGTGGGACAGAGGCC 282
QY 61 ValGluValTrpGlnGlyLeuAlaLeuLeuSerGlnAlaValLeuArgGlyGlnAlaLeu 80
DB 283 GTAGAAAGTGTGGCAGGGCCGTGGCCCTGTCTGGAAGCTGTCTGGGGCCAGAGCCCTG 342
QY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspIysAlaValSer 100
DB 343 TTGGTCAACTCTTCCCGCGGTGGAGCCCTGTGAGTGAATGATTAAGCCGTAGT 402
QY 101 GlyLeuArgSerLeuThrTrpLeuLeuArgAlaLeuGlyAlaGlnIysGlnAlaIleSer 120
DB 403 GGCCTTCGACGCTCACACACTGTCTCGAGCTGTGGGGGCCCGAAGAAAGCCATCTCC 462
QY 121 ProProAspAlaAlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLys 140
DB 463 CCTCCAGATGCGGCGCTGAGTGTCCACCTCGAACAATCACTGCTGACACTTTCGGCAA 522
QY 141 LeuPheArgValTyrSerAsnPhelLeuArgGlyLysLeuLysLeuTyrThrGlyAla 160
DB 523 CTCTTCGAGTCTACTCAATTTCCTCCGGGAAAGCTGAAGCTTTACACAGGGAGAGCA 582
QY 161 CysArgThrGlyAspArg 166
DB 583 TGCAGACAGGGGACAGA 600

RESULT 15
ABL59289 standard; DNA: 629 BP.
XX ID ABL59289
XX AC ABL59289:
XX
XX 07-OCT-2002 (first entry)
XX
XX Nucleotide sequence of a modified human erythropoietin (Epo).
XX
XX Human; erythropoietin; Epo; glycoprotein; reticulocyte production;
XX red blood cell production; anaemia; chronic renal failure;
XX acquired immunodeficiency syndrome; AIDS; cancer; bone marrow;
XX committed erythroid progenitor; ss.
XX
XX Synthetic.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers

```

FT CDS 14..619
FT /*tag= a
FT sig_peptide 14..94
FT /*tag= b
FT /*note= "secretion signal sequence"
XX WO200249673-A2.
XX
XX
XX PD 27-JUN-2002.
XX
XX PF 08-DEC-2001; 2001WO-EP14434.
XX
XX PR 20-DEC-2000; 2000EP-0127891.
XX
XX PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX
XX PI Burg J, Engel A, Franze R, Hilger B, Schurig HE, Tischer W;
XX Wozny M;
XX
XX DR WPI: 2002-566640/60.
XX P-PSDB: ABB77901.
XX
XX PT Novel conjugate of erythropoietin glycoprotein with polyethylene
XX glycol, useful for treating diseases correlated with anaemia in chronic
XX renal failure patients and acquired immunodeficiency syndrome -
XX
XX PS Disclosure: Fig 3: 40pp: English.
XX
XX CC The present sequence encodes a modified human erythropoietin (EPO)
XX protein. The EPO was extended at the N-terminal by a proteolytic
XX cleavage site. It was used to produce conjugates of the invention. The
XX specification describes a conjugate comprising an EPO glycoprotein having
XX an N-terminal alpha-amino group, chosen from human EPO (hEPO) or its
XX analogues (where hEPO is modified by addition of 1-6 glycosylation sites
XX or a rearrangement of a glycosylation site). The glycoprotein is
XX covalently linked to a poly(ethylene glycol) group. The EPO glycoprotein
XX has in vivo biological activity of causing bone marrow cells to increase
XX production of reticulocytes and red blood cells. The conjugate increased
XX circulating half-life and plasma residence time, decreased clearance,
XX increased clinical activity in vivo, improved potency and stability, when
XX compared to unmodified EPO. The EPO conjugate is useful for preparing
XX CC medications for the treatment and prophylaxis of diseases correlated with
XX anaemia in chronic renal failure patients (CRF), acquired
XX immunodeficiency syndrome (AIDS) and for treating cancer patients
XX undergoing chemotherapy. It is also useful for treating patients by
XX stimulating the division and differentiation of committed erythroid
XX progenitors in the bone marrow.
XX
XX SO Sequence 629 BP; 122 A; 203 C; 182 G; 122 T; 0 other;

Alignment Scores:
Pred. NO.: 1,17e-83 Length: 629
Score: 825.00 Matches: 162
Percent Similarity: 97.59% Conservative: 0
Best Local Similarity: 97.59% Mismatches: 4
Query Match: 97.86% Indels: 0
DB: 24 Gaps: 0

US-09-813-775c-18 (1-166) x ABL59289 (1-629)

QY 1 AAlPProAaGleuIIIEcYsAsPserArGValLeuGIuArGTYrLeuLeuGIuAlaIys 20
DB 119 GCCCCACACAGGCTCATGTGTGACAGCGAGCTGGAGAGGTACTCTTGGAGGCCAAG 178
QY 21 GIuAlaGIuAsnIIIEThrThGlyCysAlaGIuHISCySserLeuAsnGIuAsnIIIEThr 40
DB 179 GAGGCCGAGAAATATACACAGCGGCTGTGCTGACACTGCAGCTGAATGAGAAATATCACT 238
QY 41 ValProAspThrLysValAsnPhetYrAlaTrpLysArgAsn**Ser**GInGlnAla 60
DB 239 GTCCAGACACCAAGTAAATTATGCTGAGAGAGATGAGAGTGGGGCAGCAGGCC 298
QY 61 ValGIuValTrpGInGlyLeuAlaLeuLeuSerGIuAlaValLeuArgGlyGInAlaLeu 80

DB 299 GTAGAAGTCTGGCAGGGCCCTGGCCCTGTCGGAAGCTGTCGGGGCCAGGCCCTG 358
QY 81 LeuValAsnSerSerGIuProTrpGIuProLeuGIuHISValAspLysAlaValSer 100
DB 359 TTGGTCAACTCTTCCACCGCCGCGGAGGCCCTGCAGCTGATGATGAACCCCTCACT 418
QY 101 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGIuAlaGInLysGIuAlaIIEser 120
DB 419 GGCCTTCAGAGCTCACCACCTCTGCTGCGGCTCTGGAGCCAGAGAGAGAGCCATCTCC 478
QY 121 ProProAspAlaAlaSerAlaAlaProLeuArgThrIIIEThrAlaAspThrPheArgLys 140
DB 479 CCTCCAGATGGCGCCTCAGCTGCTCACTCCGAAACAATCAGCTGACACTTTCGCAAA 538
QY 141 LeuPheArgValTrpSerAsnPhetLeuArgLysLeuLysLeuTrpTrpGlyGIuAla 160
DB 539 CTTCCAGAGTCTACTCCAAATTCTCTCCGGGGAAGCTGAAAGCTGTACACAGGGAGGCC 598
QY 161 CysArgThrGlyAspArg 166
DB 599 TGCAGACAGGGGACAGA 616

Search completed: January 8, 2003, 05:27:02
Job time : 185.334 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 8, 2003, 05:15:15 ; Search time 2221.35 Seconds
(without alignments)
2174.835 Million cell updates/sec

Title: US-09-813-775C-18
Perfect score: 843
Sequence: 1 APPRLICDSRYLERYLEAK.....NFLRGKIKLYTGACRTGDR 166

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n model -DEV=xlh
-O=/cgn2.1/USPRO/US09813775/runat_07012003_153122_23818/app_query.fasta_1.718
-DB=genEmbl -OFT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPL=0 -LOOEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=bls62 -TRANS=human4.0.cdi -LIST=45
-DOCCALIGN=200 -THR.SCORE=Pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09813775@cgn.1.1741@runat.07012003.153122_23818 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGUEURY -NEG_SCORES=0 -WAIT -IONCLOG -DEV_TIMEROUT=120
-WARN_TIMEROUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl : *
1: gb_ba : *
2: gb_hcg : *
3: gb_in : *
4: gb_om : *
5: gb_ov : *
6: gb_pat : *
7: gb_ph : *
8: gb_pl : *
9: gb_pr : *
10: gb_ro : *
11: gb_sts : *
12: gb_sy : *
13: gb_un : *
14: gb_vl : *
15: em_ba : *
16: em_fun : *
17: em_hum : *
18: em_in : *
19: em_mu : *
20: em_om : *
21: em_or : *
22: em_ov : *
23: em_pat : *
24: em_ph : *
25: em_pl : *
26: em_ro : *
27: em_sts : *
28: em_un : *

29: em_vl : *
30: em_hcg_hum : *
31: em_hcg_inv : *
32: em_hcg_other : *
33: em_hcg_mus : *
34: em_hcg_pin : *
35: em_hcg_rod : *
36: em_hcg_mam : *
37: em_hcg_vrt : *
38: em_sy : *
39: em_hcgo_hum : *
40: em_hcgo_mus : *
41: em_hcgo_other : *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	825	97.9	501	6	BD007423
2	825	97.9	508	6	BD007423 Cyclical1
3	825	97.9	514	6	AX150264
4	825	97.9	533	6	AX150264 Sequence 69
5	825	97.9	551	6	I07891 Sequence 5
6	825	97.9	551	6	I00283 Sequence 4
7	825	97.9	551	6	I07892 Sequence 6
8	825	97.9	582	6	AX025442
9	825	97.9	582	6	AX025442 Sequence 6
10	825	97.9	585	6	AX046870
11	825	97.9	724	6	I05397
12	825	97.9	788	6	AR086679
13	825	97.9	823	6	AR063255
14	825	97.9	823	6	AR179013
15	825	97.9	1011	6	A47505
16	825	97.9	3100	6	AX451645
17	817	96.9	1342	9	HSEPR
18	816	96.8	501	6	AX464420
19	815	96.7	501	6	AX464421
20	802	95.1	945	6	A47500
21	802	95.1	969	6	A47497
22	743.5	88.2	1344	6	I08348
23	743.5	88.2	1462	9	MACEPO
24	743.5	88.2	5107	6	A92667
25	743.5	88.2	5107	6	AR158347
26	741.5	88.0	691	9	MACEPTYHRO
27	737.5	87.5	1344	6	I07889
28	733	87.0	1292	6	E00630
29	709	84.1	512	6	BD007363
30	703	83.4	512	6	BD007364
31	699	82.9	512	6	BD007365
32	695	82.4	681	4	CATERYTHRO
33	694	82.3	512	6	BD007366
34	692	82.1	679	10	RATEPYTHRO
35	692	82.1	1395	10	RATEPO
36	689	81.7	512	6	BD007367
37	688	81.6	725	4	FD000685
38	683	81.0	512	6	BD007368
39	679.5	80.6	1133	4	BTU44762
40	679.5	80.6	1258	4	BOVEPO
41	676	80.2	3609	6	AX249944
42	675.5	80.1	588	4	AP290943
43	675	80.1	681	4	DOSEPTIRE
44	674	80.0	512	6	BD007369
45	672.5	79.8	585	4	OAEERYPOIA

RESULT 1

ALIGNMENTS

BD007423 501 bp DNA linear PAT 31-JAN-2002
LOCUS BD007423
DEFINITION Cyclically rearranged erythropoietin receptor agonist.
ACCESSION BD007423
VERSION BD007423.1 GI:18635794
KEYWORDS JP 2001503266-A/61.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 501)
AUTHORS McWhorter,C.A., Feng,Y. and Somers,N.
TITLE Cyclically rearranged erythropoietin receptor agonist
JOURNAL Patent: JP 2001503266-A 61 13-MAR-2001;
G.D. SEARLE & CO
COMMENT OS Unidentified
PN JP 2001503266-A/61
PD 13-MAR-2001
PF 23-OCT-1997 JP 1998520528
PI 25-OCT-1996 US 60/034044
PC CHARLES A MCWHERTER,YI QING FENG,NINA SOMERS
PC C12N15/09,A61K35/18,A61K35/28,A61K38/22,A61K48/00,A61P7/06,PC
A61P43/00,
PC C07K14/52,C12N5/06,C12P21/02,C12N15/00,C12N5/00,A61K37/24 CC
Strandedness: Single;
CC Topology: Linear;
FT key Location/Qualifiers
FT source 1..501 /organism='unidentified',
location/Qualifiers
1..501
BASE COUNT 108 a 152 c 145 g 96 t
ORIGIN
Alignment Scores:
Pred. No.: 1.07e-77 Length: 501
Score: 825.00 Matches: 162
Percent Similarity: 97.59% Conservative: 0
Best Local Similarity: 97.59% Mismatches: 4
Query Match: 97.86% Indels: 0
Gaps: 0
DB: 6
US-09-813-775c-18 (1-166) x BD007423 (1-501)
QY 1 AAlaProPArgLeuIleCysAspSerArgValLeuGluArgTyrLeuLeuGluAlaLys 20
Db 1 GCCCACCACGCGTCATCTGTGACAGCCGAGTCTCGAGAGTACTCTTGAGGCCAAG 60
QY 21 GluAlaGluAsnIleThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
Db 61 GAGCGCGGAATATACACAGCGGCTGTCTCAACACTGCAAGTGAATGAGAAATATCACT 120
QY 41 ValProAspThrIleValAsnPhetYrAlaTrpLysArgAsn***Ser**GlnGlnAla 60
Db 121 GTCCAGACACCAAAAGTTATTTCTATGCTCGGAAGAGATGAGAGTGGCGGACAGCGCC 180
QY 61 ValGluValTrpGlnGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
Db 181 GTAGAAAGTGTGGAGAGGCTGCTGCTGTGCGAAAGCTGTCTCGGGGCGACGCCCTCG 240
QY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSer 100
Db 241 TTGGTCAACTCTTCCACCGCTGGAGGCCCTGCGAGCTGCTGGAATGGAAGCCGTCAGT 300
QY 101 GlyLeuArgSerLeuThrThLeuLeuArgAlaLeuGlyAlaGlnLysGluAlaIleSer 120
Db 301 GGCCTTCGACGCTCCACACTCTGCTTGGCTCGGAGGCCACAAAGAGCCATCTCC 360
QY 121 ProProAspAlaAlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLys 140
Db 361 CCTCAGATGGGGCTCAGCTGCTCCACTCGGAACATCATCTGTGACACTTTCCGCAA 420

QY 141 LeuPheArgValTyrSerAsnPhetLeuArgGlyLysLeuLysLeuTyrThrGlyGluAla 160
Db 421 CTTCTCCAGAGTCTACTCCAAATTTCTCCGGGGAAGCTGAAGCTGTACACAGGAGGCC 480
QY 161 CysArgThrGlyAspArg 166
Db 481 TGCAGACAGGGACAGA 498
RESULT 2
LOCUS 149875 508 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 69 from patent US 5641663.
ACCESSION 149875
VERSION 149875.1 GI:2472095
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 508)
AUTHORS Garvin,R.T. and Malek,L.T.
TITLE Expression system for the secretion of bioactive human granulocyte
macrophage colony stimulating factor (GM-CSF) and other
heterologous proteins from streptomyces
JOURNAL Patent: US 5641663-A 69 24-JUN-1997;
FEATURES
source 1..508 /organism='unknown',
location/Qualifiers
1..508
BASE COUNT 82 a 193 c 161 g 72 t
ORIGIN
Alignment Scores:
Pred. No.: 1.09e-77 Length: 508
Score: 825.00 Matches: 162
Percent Similarity: 97.59% Conservative: 0
Best Local Similarity: 97.59% Mismatches: 4
Query Match: 97.86% Indels: 0
Gaps: 0
DB: 6
US-09-813-775c-18 (1-166) x 149875 (1-508)
QY 1 AAlaProPArgLeuIleCysAspSerArgValLeuGluArgTyrLeuLeuGluAlaLys 20
Db 5 GCCCCCCGCGCTCATCTGCGACAGCCGCGTCTCGAGCGGTACTCTCGAAGCCAAAG 64
QY 21 GluAlaGluAsnIleThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
Db 65 GAGCGGAGAAATATACACAGCGGAGTGGCCGCGACACTGCTCCACAGAGATACAC 124
QY 41 ValProAspThrIleValAsnPhetYrAlaTrpLysArgAsn***Ser**GlnGlnAla 60
Db 125 GTCCCGACACCAAGTCAACTTCTACCGCTGCAAGCCACAGAGTGGCGGACAGCGG 184
QY 61 ValGluValTrpGlnGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
Db 185 GTCGAGTGTGGAGAGGCTCGCGTCTCTCCAGAGGCTCTCCGGGCGACGCCCTC 244
QY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSer 100
Db 245 CTGGTGAACGTGTCACACCGCGGAGCGCGCTCCAGCTGCAAGTGCAGACAGCGCTCC 304
QY 101 GlyLeuArgSerLeuThrThLeuLeuArgAlaLeuGlyAlaGlnLysGluAlaIleSer 120
Db 305 GGCCTCCGCTCCGACACCGCTGCTGCGGCCCTCGGTCGCGCAAGAGAGGCCATCTCG 364
QY 121 ProProAspAlaAlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLys 140
Db 365 CCCCCGAGCGCGCGAGCGCGCGCGCTGCGAGATCAGCGGAGACACTTCGCAAG 424
QY 141 LeuPheArgValTyrSerAsnPhetLeuArgGlyLysLeuLysLeuTyrThrGlyGluAla 160
Db 425 CTGTTCCGGGCTCTACTGCACTTCTCTCGGGGGAAGCTGAAGCTTCTACACCGGAGGCC 484
QY 161 CysArgThrGlyAspArg 166

```

|||||
Db      485  TGCCGACGCGGACCGC  502
RESULT 3
AX150264
LOCUS      AX150264          514 bp      DNA      linear      PAT 08-JUN-2001
DEFINITION Sequence 1 from Patent WO0136489.
ACCESSION  AX150264
VERSION     AX150264.1  GI:14348284
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE   1 (bases 1 to 514)
AUTHORS     Hartmann,A., Brandt,S., Rieke,E., Sobel,C., Lo,K.M., Way,J.C. and
            Gillies,S.
            Erythropoietin forms with improved properties
            Patent: WO 0136489-A 1 25-MAY-2001;
            MERCK PATENT GmbH (DE)
FEATURES
source      1..514
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            <8..508
            /note="Human EPO, DNA sequence modified but no change in
            protein sequence"
            /protein_id="CAC41224.1"
            /db_xref="GI:14348285"
            /translation="APRLICDSRVLEERYLEAKAEINITTCARHCISINENITVPDT
            /KVFYAKRMVEGOAOVEWOGALLSEAVLWGQALVNSOPWEPLOLHYDKAVSGL
            RSLTTLRLALGAKKEAISPDDAASAPRITITADIFRKLFRVSNELRGLKLTLYGEA
            CRTGDR"
BASE COUNT  106 a 159 c 152 g 97 t
ORIGIN
Alignment Scores:
Pred. No.: 1,11e-77      Length: 514
Score:      825.00      Matches: 162
Percent Similarity: 97.59%      Conservative: 0
Best Local Similarity: 97.59%      Mismatches: 4
Query Match: 97.86%      Indels: 0
            Gaps: 0
US-09-813-775C-18 (1-166) x AX150264 (1-514)
QY      1  AAlaProProArgLeuIleCysAspSerArgValLeuGluArgTyrLeuLeuGluAlaLys  20
      |||||||
Db      8  GGCACACGCGCTCATCTGTGTGACAGCCAGTCTGAGAGTACCTCTTGAGGCGCAAG  67
QY      21 GAlaAGlAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr  40
      |||||||
Db      68 GAGCGCGAGAAATATCATCAGACCGCTGTCTGAACTGCGAGCTTGAAATGAGAACTATCACC  127
QY      41 ValProAspThrLysValAsnPhetYrAlaTPrLysArgAsn**Ser**GlnGlnAla  60
      |||||||
Db      128 GTGGCTGACACCAAGTGAATTTCTATGCTTGAGAGATGAGATTGGCGGCGCAGGCGC  187
QY      61 ValGluValTPrGlnGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu  80
      |||||||
Db      188 GTAGAGAGTGTGGAGGCGCTGCGCTGTCTGCGAAGCTGTCTGCGGCGCAGGCGCTG  247
QY      81 LeuValAsnSerSerGlnProTPrGluProLeuGlnLeuHisValAspLysAlaValSer  100
      |||||||
Db      248 TTGGTGAACCTCTTCCAGCGCTGTGAGAGCCCTTCGCAACTGATGTGATTAAGCGCTGAGT  307
QY      101 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGluAlaIleSer  120
      |||||||
Db      308 GGGCTTCGAGCGCTCAACACTGTGCTTGGGCTGTGGAGCCGAGAAAGGCAATCTCC  367
QY      121 ProProAspAlaIleAsnAlaIleProLeuArgThrIleThrAlaAspThrPheArgLys  140
      |||||||

```

```

Db      368 CATTCAAGATCCGGCTTCAGCTGCTCCCTCCGACAAATCATCTGACACTTCCGCAAA  427
QY      141 LeuPheArgValTyrSerAsnPhelLeuArgGlyLysLeuLysLeuTyrThrGlyGluAla  160
      |||||||
Db      428 CTCTTCGAGATCTACTCCAAATTTCTTCGCGGAGAAAGCTGAAGCTGTACACAGGGAGGCC  487
QY      161 CysArgThrGlyAspArg 166
      |||||||
Db      488 TGCCGAGACAGGGAGACAGA  505
RESULT 4
107891
LOCUS      107891          533 bp      DNA      linear      PAT 02-DEC-1994
DEFINITION Sequence 5 from Patent EP 0148605.
ACCESSION  107891
VERSION     107891.1  GI:589396
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 533)
AUTHORS     Lin,F.-K.
            Production of erythropoietin
            Patent: EP 0148605-A2 5 17-JUL-1985;
            Location/Qualifiers
FEATURES
source      1..533
            /organism="unknown"
BASE COUNT  131 a 134 c 138 g 130 t
ORIGIN
Alignment Scores:
Pred. No.: 1.16e-77      Length: 533
Score:      825.00      Matches: 162
Percent Similarity: 97.59%      Conservative: 0
Best Local Similarity: 97.59%      Mismatches: 4
Query Match: 97.86%      Indels: 0
            Gaps: 0
US-09-813-775C-18 (1-166) x 107891 (1-533)
QY      1  AAlaProProArgLeuIleCysAspSerArgValLeuGluArgTyrLeuLeuGluAlaLys  20
      |||||||
Db      30  GCTCCGCGCGCTGTGATCTGCACTCGAGAGTTCTGCAACGTTACTGCTGGAAGCTAAA  89
QY      21 GAlaAGlAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr  40
      |||||||
Db      90 GAAGCTGAAGAAACATCAACCACTGCTGTGCTGAACTGTTTGAACGAAACATTAACG  149
QY      41 ValProAspThrLysValAsnPhetYrAlaTPrLysArgAsn**Ser**GlnGlnAla  60
      |||||||
Db      150 GTACCAAGACACCAAGTTACTTCTACGCTTGGAAGCGTATGGAAGTTGGTCAACAAGCA  209
QY      61 ValGluValTPrGlnGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu  80
      |||||||
Db      210 GTTGAAGTTTGGCAGGAGTGTGGCACTGCTGACGAGGCTGTACTGCTGCGCAGGCACTG  269
QY      81 LeuValAsnSerSerGlnProTPrGluProLeuGlnLeuHisValAspLysAlaValSer  100
      |||||||
Db      270 CTGGTAAACCTCTCTACGCGCTGGGAACCGCTGAGCTGATGTGACAAAGCAAGTATCT  329
QY      101 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGluAlaIleSer  120
      |||||||
Db      330 GGGCTGAGATCTGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT  389
QY      121 ProProAspAlaIleAsnAlaIleProLeuArgThrIleThrAlaAspThrPheArgLys  140
      |||||||
Db      390 CCGCGGAGTCTGCATCTGCTGCGACCGCTGCGAATCACTGCTGATTAACCTTCGCGAAA  449
QY      141 LeuPheArgValTyrSerAsnPhelLeuArgGlyLysLeuLysLeuTyrThrGlyGluAla  160
      |||||||
Db      450 CTGTTTGTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT  509
QY      161 CysArgThrGlyAspArg 166

```

DB 510 TCCGCTACTGGTACCGC 527
RESULT 5
LOCUS 100283
DEFINITION Sequence 4 from Patent US 4935350.
ACCESSION 100283
VERSION 100283.1
KEYWORDS GI:314048
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 551)
AUTHORS Patel,A.C. and Ratzkin,B.J.
TITLE Materials and methods for controlling plasmid copy number and stability
JOURNAL Patent: US 4935350-A 4-19-JUN-1990;
COMMENT On Jul 30, 1993 this sequence version replaced gi:285577.
FEATURES
source Location/Qualifiers
BASE COUNT 157 a 112 c 129 g 153 t
ORIGIN
Alignment Scores:
Pred. NO.: 1.2e-77 Length: 551
Score: 825.00 Matches: 162
Percent Similarity: 97.59% Conservative: 0
Best Local Similarity: 97.59% Mismatches: 4
Query Match: 97.86% Indels: 0
DB: Gaps: 6
US-09-813-775C-18 (1-166) x 100283 (1-551)
QY 1 AAlaProProAlaGLeuILllecysAspSerArGValLeuGluArGTYrLeuLeuGluAlaLys 20
DB 16 GCTCCACCACCAATGATGTGCTACTCGAGAGTTTGGAAAGATCTTGTGGAAAGCTTAA 75
QY 21 GluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
DB 76 GAAGCTGAAACATCACCACCTGGTGTGCTGACACCTGTTCTTGAAGAAACCTTACG 135
QY 41 ValProAspThrLysValAsnPhetYrAlaTrpLysArGAsn**Ser**GlnGlnAla 60
DB 136 GTACCAGACACCAGGTTAACTTACCTTGAACCTGATGAGAGTTGTCACACAGCT 195
QY 61 ValGluValITrPGInGlyLeuAlaLeuLeuSerGluAlaValLeuArGlyGlnAlaLeu 80
DB 196 GTTGAAGTTTGGCAAGTTTGGCTTGTATCTGAAGCTGTTTGAAGGTCACAGCTTGG 255
QY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSer 100
DB 256 TTGCTTAACCTCTTCCACACCATGGSAACCATTTGCAATTGCAAGTGAATTAAGCCGTCT 315
QY 101 GlyLeuArGSerLeuThrThrLeuLeuArGAlaLeuGlyAlaGlnLysGlnAlaIleSer 120
DB 316 GGTTCAGATCTTTCACACTTGTGAGACCTTGGGTCGCAAAAGAACCCATTTC 375
QY 121 ProProAspAlaAlaSerAlaAlaProLeuArGThrIleThrAlaAspThrPhaArgLys 140
DB 376 CCACCAGACGCTTTCGCGCTCCATTTGAGAACCATCAGCTGATACCTTCAGAAAG 435
QY 141 LeuPheArGValITySerAsnPhaLeuArGlyLysLeuLysLeuTYrThrGlyGlnAla 160
DB 436 TTATTCAAGATTACTCCAACTTCTTGAGAGGTAATTGAAGTTTATACACCGGTAAAGCC 495
QY 161 CysArGThrGlyAspArg 166
DB 496 TGTAGACTGCTGACAGA 513
RESULT 6

107892
LOCUS 107892 551 bp DNA Linear PAT 02-DEC-1994
DEFINITION Sequence 6 from Patent EP 0148605.
ACCESSION 107892
VERSION 107892.1
KEYWORDS GI:589397
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 551)
AUTHORS Llo,F.-K.
TITLE Production of erythropoietin
JOURNAL Patent: EP 0148605-A2 6-17-JUL-1985;
FEATURES
source Location/Qualifiers
BASE COUNT 157 a 112 c 129 g 153 t
ORIGIN
Alignment Scores:
Pred. NO.: 1.2e-77 Length: 551
Score: 825.00 Matches: 162
Percent Similarity: 97.59% Conservative: 0
Best Local Similarity: 97.59% Mismatches: 4
Query Match: 97.86% Indels: 0
DB: Gaps: 6
US-09-813-775C-18 (1-166) x 107892 (1-551)
QY 1 AAlaProProAlaGLeuILllecysAspSerArGValLeuGluArGTYrLeuLeuGluAlaLys 20
DB 16 GCTCCACCACCAATGATGTGCTACTCGAGAGTTTGGAAAGATCTTGTGGAAAGCTTAA 75
QY 21 GluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
DB 76 GAAGCTGAAACATCACCACCTGGTGTGCTGACACCTGTTCTTGAAGAAACCTTACG 135
QY 41 ValProAspThrLysValAsnPhetYrAlaTrpLysArGAsn**Ser**GlnGlnAla 60
DB 136 GTACCAGACACCAGGTTAACTTACCTTGAACCTGATGAGAGTTGTCACACAGCT 195
QY 61 ValGluValITrPGInGlyLeuAlaLeuLeuSerGluAlaValLeuArGlyGlnAlaLeu 80
DB 196 GTTGAAGTTTGGCAAGTTTGGCTTGTATCTGAAGCTGTTTGAAGGTCACAGCTTGG 255
QY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSer 100
DB 256 TTGCTTAACCTCTTCCACACCATGGSAACCATTTGCAATTGCAAGTGAATTAAGCCGTCT 315
QY 101 GlyLeuArGSerLeuThrThrLeuLeuArGAlaLeuGlyAlaGlnLysGlnAlaIleSer 120
DB 316 GGTTCAGATCTTTCACACTTGTGAGACCTTGGGTCGCAAAAGAACCCATTTC 375
QY 121 ProProAspAlaAlaSerAlaAlaProLeuArGThrIleThrAlaAspThrPhaArgLys 140
DB 376 CCACCAGACGCTTTCGCGCTCCATTTGAGAACCATCAGCTGATACCTTCAGAAAG 435
QY 141 LeuPheArGValITySerAsnPhaLeuArGlyLysLeuLysLeuTYrThrGlyGlnAla 160
DB 436 TTATTCAAGATTACTCCAACTTCTTGAGAGGTAATTGAAGTTTATACACCGGTAAAGCC 495
QY 161 CysArGThrGlyAspArg 166
DB 496 TGTAGACTGCTGACAGA 513
RESULT 7
LOCUS 108350 551 bp DNA Linear PAT 02-DEC-1994
DEFINITION Sequence 6 from Patent WO 8502610.
ACCESSION 108350
VERSION 108350.1
KEYWORDS GI:588940
SOURCE Unknown.

TITLE Recombinant human erythropoietin
JOURNAL Patent: EP 0267678-A1 2 18-MAY-1988;
FEATURES Location/Qualifiers
source 1..582
/organism="unknown"

BASE COUNT 113 a 181 c 170 g 118 t
ORIGIN

Alignment Scores:

Pred. No.: 1.28e-77 Length: 582
Score: 825.00 Matches: 162
Percent Similarity: 97.59% Conservative: 0
Best Local Similarity: 97.59% Mismatches: 4
Query Match: 97.86% Indels: 0
DB: Gaps: 0

US-09-813-775c-18 (1-166) x I05397 (1-582)

QY 1 AAlProProArgLeuIleCysAspSerArgValLeuGluArgTyrLeuLeuGluAlaLys 20
DB 82 GCGCCACACCGCTCATCTGTGACAGCCGAGTCTGGAGAGGTACCTCTGGAGGCCAAG 141
QY 21 GluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
DB 142 GAGCCGGAATATATCAGCAGGGGCTGTGCTGAACACTGCACTGATGAGATATCACT 201
QY 41 ValProAspThrLysValAsnPhetYrAlaTrpLysArgAsn***Ser**GlnGlnAla 60
DB 202 GTCCAGACACCAAGTTAAATTCTATGCTGAGAGAGATGAGAGTGGCGACAGCGCC 261
QY 61 ValGluValTrpGlnGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
DB 262 GTAGAGAGTGTGAGAGGCTGGCCCTGCTGTGGAAGTGTCTGCGGGCCAGCGCCCTG 321
QY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSer 100
DB 322 TTGGTCAACTCTTCCAGCGCTGGAGGCCCTGCACTGATGATTAAGCCGTCACT 381
QY 101 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGluAlaIleSer 120
DB 382 GGCCTTCGAGGCTCACCACCTGCTGCTGGGAGCCAGAGGAGGAGGAGGAGGAGGAGG 441
QY 121 ProProAspAlaIleSerAlaIleProLeuArgThrIleThrAlaAspThrPheArgLys 140
DB 442 CCTCAGATGCGGCTCAGCTGCTCAGCAGCAATCATCTGCTGCTGCTGCTGCTGCTG 501
QY 141 LeuPheArgValTyrSerAsnPhetLeuArgGlyLysLeuLysLeuTyrThrGlyGluAla 160
DB 502 CTTTCCGAGTCTACTCCAAATTCTCTCGGGGAAAGCTGAGCTGATACAGAGGAGGCC 561
QY 161 CysArgThrGlyAspArg 166
DB 562 TGCAGGACAGGGGACAGA 579

RESULT 10

AX046870 AX046870 585 bp DNA linear PAT 15-DEC-2000
LOCUS Sequence 3 from Patent WO0068376.
DEFINITION AX046870
ACCESSION AX046870
VERSION AX046870.1 GI:11876340
KEYWORDS

SOURCE

ORGANISM Chimpanzee.
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
REFERENCE 1 (bases 1 to 585)
AUTHORS Desauvage, F. and Henner, D.J.
TITLE Chimpanzee erythropoietin (chepo) polypeptides and nucleic acids
JOURNAL encoding the same
GENENTECH, INC. (US)
Patent: WO 0068376-A 3 16-NOV-2000;
Location/Qualifiers

FEATURES
Source 1..585

/organism="Pan troglodytes"
/db_xref="taxon:9598"
BASE COUNT 115 a 185 c 168 g 117 t
ORIGIN

Alignment Scores:

Pred. No.: 1.29e-77 Length: 585
Score: 825.00 Matches: 162
Percent Similarity: 97.59% Conservative: 0
Best Local Similarity: 97.59% Mismatches: 4
Query Match: 97.86% Indels: 0
DB: Gaps: 0

US-09-813-775c-18 (1-166) x AX046870 (1-585)

QY 1 AAlProProArgLeuIleCysAspSerArgValLeuGluArgTyrLeuLeuGluAlaLys 20
DB 82 GCGCCACACCGCTCATCTGTGACAGCCGAGTCTGGAGAGGTACCTCTGGAGGCCAAG 141
QY 21 GluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
DB 142 GAGCCGGAATATATCAGCAGGGGCTGTGCTGAACACTGCACTGATGAGATATCACT 201
QY 41 ValProAspThrLysValAsnPhetYrAlaTrpLysArgAsn***Ser**GlnGlnAla 60
DB 202 GTCCAGACACCAAGTTAAATTCTATGCTGAGAGAGATGAGAGTGGCGACAGCGCC 261
QY 61 ValGluValTrpGlnGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
DB 262 GTAGAGAGTGTGAGAGGCTGGCCCTGCTGTGGAAGTGTCTGCGGGCCAGCGCCCTG 321
QY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSer 100
DB 322 TTGGTCAACTCTTCCAGCGCTGGAGGCCCTGCACTGATGATTAAGCCGTCACT 381
QY 101 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGluAlaIleSer 120
DB 382 GGCCTTCGAGGCTCACCACCTGCTGCTGGGAGCCAGAGGAGGAGGAGGAGGAGGAGG 441
QY 121 ProProAspAlaIleSerAlaIleProLeuArgThrIleThrAlaAspThrPheArgLys 140
DB 442 CCTCAGATGCGGCTCAGCTGCTCAGCAGCAATCATCTGCTGCTGCTGCTGCTGCTG 501
QY 141 LeuPheArgValTyrSerAsnPhetLeuArgGlyLysLeuLysLeuTyrThrGlyGluAla 160
DB 502 CTTTCCGAGTCTACTCCAAATTCTCTCGGGGAAAGCTGAGCTGATACAGAGGAGGCC 561
QY 161 CysArgThrGlyAspArg 166
DB 562 TGCAGGACAGGGGACAGA 579

RESULT 11

I05399 I05399 724 bp DNA linear PAT 02-DEC-1994
LOCUS Sequence 4 from Patent EP 0267678.
DEFINITION I05399
ACCESSION I05399
VERSION I05399.1 GI:590971
KEYWORDS

SOURCE

ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 724)
AUTHORS Beck, A.K., Wilthy, R.M., Zabrecky, J.R. and Mastello, N.C.
TITLE Recombinant human erythropoietin
JOURNAL Patent: EP 0267678-A1 4 18-MAY-1988;
FEATURES Location/Qualifiers
source 1..724
/organism="unknown"

BASE COUNT 141 a 235 c 202 g 146 t
ORIGIN

Alignment Scores:

Pred. No.: 1.65e-77 Length: 724

Score: 825.00 Matches: 162
 Best Local Similarity: 97.59% Conservative: 0
 Best Local Similarity: 97.59% Mismatches: 4
 Query Match: 97.86% Indels: 0
 DB: 6 Gaps: 0

US-09-813-775C-18 (1-166) x I05399 (1-724)

QY 1 AAlaProArgLeuIleCysAspSerArgValLeuGluArgTyrLeuGluAlaLys 20
 DB 131 GCCCCACACCGCTCATCTGTGACAGCCAGTCCGTGAGAGTACTCTTGGAGGCCAAG 190
 QY 21 GluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
 DB 191 GAGCCGGAATAATCATCAGCAGCGGCTGTGGAACACTGAGCTTGAATGAGATATACACT 250
 QY 41 ValProAspThrLysValAsnPhetYrAlaTrrPlyArgAsn***Ser***GlnGlnAla 60
 DB 251 GTCCCGACACCAAAAGTTAATTTCTATGCTTGAGAGAGATGAGGTCCGGGACGAGGCC 310
 QY 61 ValGluValTrrPdlngIlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
 DB 311 GTAGAAGTCTGGAGAGGCTCGGCCCTGCTGTGGAAGCTGTCTCGGGGGCCAGGCCCTG 370
 QY 81 LeuValAsnSerSerGlnProTrrPgluProLeuGlnLeuHisValAspLysAlaValSer 100
 DB 371 TTGGTCAACTCTTCCACGCGCTGGAGCCCTGCACCTGCATGCGATTAAGCCGTCAGT 430
 QY 101 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGluAlaIleSer 120
 DB 431 GGCCTTCGACGCTCACCACTGCTTGCGGCTCTGGAGCCAGAGAGGAGCCATCTCC 490
 QY 121 ProProAspAlaAlaSerAlaAlaProLeuArgTrrIleThrAlaAspThrPheArgLys 140
 DB 491 CCTCCAGATGCGGCGCTCACCTGCTGCACGCCAGAACATCATGCTGACACTTCCGCAAA 550
 QY 141 LeuPheArgValTyrSerAsnPhetLeuArgGlyLysLeuLysLeuTyrThrGlyAla 160
 DB 551 CTTCTCCGAGTCTACCAATTTCTCCGCGGAAAGCTGAAGCTGTACACAGGAGGCC 610
 QY 161 CysArgThrGlyAspArg 166
 DB 611 TGCAGGACAGGGGACAGA 628

RESULT 12

LOCUS AR086679 788 bp DNA linear PAT 07-SEP-2000
 DEFINITION Sequence 35 from patent US 5985607.
 ACCESSION AR086679
 VERSION AR086679.1 GI:10013445
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE

1 (bases 1 to 788)
 DeJucue, G. and Awang, G.
 Recombinant DNA molecules and expression vectors for tissue
 plasminogen activator
 Patent: US 5985607-A 35 16-NOV-1999;
 JOURNAL Location/Qualifiers
 FEATURES
 source 1..788
 BASE COUNT 145 a 264 c 228 g 151 t
 ORIGIN

Alignment Scores:

Pred. No.: 1,82e-77 Length: 788
 Score: 825.00 Matches: 162
 Percent Similarity: 97.59% Conservative: 0
 Best Local Similarity: 97.59% Mismatches: 4
 Query Match: 97.86% Indels: 0
 DB: 6 Gaps: 0

US-09-813-775C-18 (1-166) x AR086679 (1-788)

QY 1 AAlaProArgLeuIleCysAspSerArgValLeuGluArgTyrLeuGluAlaLys 20
 DB 161 GCCCCACACCGCTCATCTGTGACAGCCAGTCCGTGAGAGTACTCTTGGAGGCCAAG 220
 QY 21 GluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 40
 DB 221 GAGCCGGAATAATCATCAGCAGCGGCTGTGGAACACTTGAAGCTTGAATGAGATATACACT 280
 QY 41 ValProAspThrLysValAsnPhetYrAlaTrrPlyArgAsn***Ser***GlnGlnAla 60
 DB 281 GTCCCGACACCAAAAGTTAATTTCTATGCTTGAGAGAGATGAGGTCCGGGACGAGGCC 340
 QY 61 ValGluValTrrPdlngIlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeu 80
 DB 341 GTAGAAGTCTGGAGAGGCTCGGCCCTGCTGTGGAAGCTGTCTCGGGGGCCAGGCCCTG 400
 QY 81 LeuValAsnSerSerGlnProTrrPgluProLeuGlnLeuHisValAspLysAlaValSer 100
 DB 401 TTGGTCAACTCTTCCACGCGCTGGAGCCCTGCACCTGCATGCGATTAAGCCGTCAGT 460
 QY 101 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGluAlaIleSer 120
 DB 461 GGCCTTCGACGCTCACCACTGCTTGCGGCTCTGGAGCCAGAGAGGAGCCATCTCC 520
 QY 121 ProProAspAlaAlaSerAlaAlaProLeuArgTrrIleThrAlaAspThrPheArgLys 140
 DB 521 CCTCCAGATGCGGCGCTCACCTGCTGCACGCCAGAACATCATGCTGACACTTCCGCAAA 580
 QY 141 LeuPheArgValTyrSerAsnPhetLeuArgGlyLysLeuLysLeuTyrThrGlyAla 160
 DB 581 CTTCTCCGAGTCTACCAATTTCTCCGCGGAAAGCTGAAGCTGTACACAGGAGGCC 640
 QY 161 CysArgThrGlyAspArg 166
 DB 641 TGCAGGACAGGGGACAGA 658

RESULT 13

LOCUS AR063255 823 bp DNA linear PAT 29-SEP-1999
 DEFINITION Sequence 1 from patent US 5846528.
 ACCESSION AR063255
 VERSION AR063255.1 GI:5992563
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE

1 (bases 1 to 823)
 Podsakoff, G.M. and Kurtzman, G.J.
 Treating anemia using recombinant adeno-associated virus virions
 comprising an Epo DNA sequence
 Patent: US 5846528-A 1 08-DEC-1999;
 JOURNAL Location/Qualifiers
 FEATURES
 source 1..823
 BASE COUNT 159 a 275 c 233 g 156 t
 ORIGIN

Alignment Scores:

Pred. No.: 1,92e-77 Length: 823
 Score: 825.00 Matches: 162
 Percent Similarity: 97.59% Conservative: 0
 Best Local Similarity: 97.59% Mismatches: 4
 Query Match: 97.86% Indels: 0
 DB: 6 Gaps: 0

US-09-813-775C-18 (1-166) x AR063255 (1-823)

QY 1 AAlaProArgLeuIleCysAspSerArgValLeuGluArgTyrLeuGluAlaLys 20
 DB 133 GCCCCACACCGCTCATCTGTGACAGCCAGTCCGTGAGAGTACTCTTGGAGGCCAAG 192


```
Db 262 GTAGAGTCTGGCAGGGCTGGCCCTGCTGCGAAGCTGTCTGGCGGGCCAGGCCCTG 321
QY 81 LeuValAsnSerSerGlnProTrpGluProLeuGlnIleuHisValAspLysAlaValSer 100
Db 322 TTGGTCAACTCTTCCCGCGGTGGAGCCCTGCGAGCTGCATGTGGATTAAGCCGTCAGT 381
QY 101 GlyLeuArgSerLeuThrLeuLeuArgAlaLeuGlyAlaGlnLysGluAlaIleSer 120
Db 382 GGCCCTTCGCGAGCTTCACACCTGCTGCGGGCTCTGGGAGCCCGAAGGAGCCATCTCC 441
QY 121 ProProAspAlaAlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLys 140
Db 442 CCTCCAGATCGGCTCAGCTGCTCCACTCGAACAATCACTGCTGACACTTTCGGCAA 501
QY 141 LeuPheArgValTyrSerAsnPheLeuArgGlyLysLeuLysLeuTyrThrGlyGluAla 160
Db 502 CTCTTCGAGTCTACTCCAAATTTCCTCCGGGAAAGCTGAAGCTGTACACAGGGAGGCC 561
QY 161 CysArgThrGlyAspArg 166
Db 562 TGCAGGACAGGGAGCAGA 579
```

Search completed: January 8, 2003, 06:57:33
Job time : 2227.35 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 03:41:38 ; Search time 28.6685 Seconds
(without alignments)
112.338 Million cell updates/sec

Title: US-09-813-775c-18

Perfect score: 843

Sequence: 1 APRRLICDSRYLERYLLLEAK.....NFLRGKILTYGECRCRIGDR 166

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 118974 seqs, 19401057 residues

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications, AA.*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	825	97.9	166	10	US-09-853-731-2
2	825	97.9	166	12	US-10-014-363-2
3	825	97.9	169	12	US-10-014-363-4
4	825	97.9	174	12	US-10-014-363-3
5	825	97.9	174	12	US-10-014-363-5
6	820	97.3	165	10	US-09-853-731-1
7	820	97.3	165	12	US-10-014-363-1
8	284	33.7	60	10	US-09-864-761-48502
9	111	13.2	30	10	US-09-975-063-2
10	111	13.2	30	12	US-10-011-858-2
11	108	12.8	20	9	US-09-919-703-9
12	88	10.4	20	9	US-09-919-703-10
13	76.5	9.1	2472	10	US-09-815-242-5064
14	69.5	8.2	542	9	US-10-045-815-8
15	69.5	8.2	747	9	US-10-045-815-6
16	69.5	8.2	813	9	US-09-964-899-25
17	68	8.1	321	9	US-09-945-182-26
18	67.5	8.0	1564	10	US-09-801-368-244
19	66.5	7.9	119	9	US-09-738-626-6842

20	66.5	7.9	766	10	US-09-925-301-1276	Sequence 1276, Ap
21	66	7.8	210	10	US-09-731-872-387	Sequence 387, App
22	66	7.8	409	10	US-09-815-242-14050	Sequence 14050, A
23	66	7.8	7257	9	US-10-014-717-5	Sequence 5, Appl1
24	65.5	7.8	240	9	US-09-738-626-6760	Sequence 6760, Ap
25	65.5	7.8	711	9	US-09-738-626-3507	Sequence 3507, Ap
26	64.5	7.7	379	9	US-09-975-139-8	Sequence 8, Appl1
27	64.5	7.7	552	10	US-09-880-192-50	Sequence 50, Appl1
28	64.5	7.7	2504	10	US-09-817-514A-8	Sequence 8, Appl1
29	64	7.6	538	9	US-09-738-626-4165	Sequence 4165, Ap
30	64	7.6	954	9	US-09-944-413-7	Sequence 7, Appl1
31	64	7.6	954	9	US-09-944-403-7	Sequence 7, Appl1
32	64	7.6	954	9	US-09-944-896-7	Sequence 7, Appl1
33	64	7.6	954	9	US-09-944-944-7	Sequence 7, Appl1
34	64	7.6	954	9	US-09-944-907-7	Sequence 7, Appl1
35	64	7.6	954	9	US-09-944-929-7	Sequence 7, Appl1
36	64	7.6	954	10	US-09-866-028-7	Sequence 7, Appl1
37	64	7.6	954	10	US-09-944-449-7	Sequence 7, Appl1
38	64	7.6	954	10	US-09-944-457-7	Sequence 7, Appl1
39	64	7.6	954	10	US-09-944-862-7	Sequence 7, Appl1
40	64	7.6	954	10	US-09-945-587-7	Sequence 7, Appl1
41	64	7.6	954	10	US-09-945-015-7	Sequence 7, Appl1
42	64	7.6	954	10	US-09-944-396-7	Sequence 7, Appl1
43	64	7.6	954	10	US-09-944-097-7	Sequence 7, Appl1
44	64	7.6	954	10	US-09-944-432-7	Sequence 7, Appl1
45	64	7.6	954	10	US-09-943-762-7	Sequence 7, Appl1

ALIGNMENTS

```

RESULT 1
US-09-853-731-2
; Sequence 2, Application US/09853731
; Patent No. US20020037841A1
; GENERAL INFORMATION:
; APPLICANT: Paradiem, Inc., Apollon
; TITLE OF INVENTION: Erythropoietin Composition
; FILE REFERENCE: 20619 US
; CURRENT APPLICATION NUMBER: US/09/853, 731
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: EP/00110355.5
; PRIOR FILING DATE: 2000-05-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-853-731-2

Query Match          97.9%; Score 825; DB 10; Length 166;
Best Local Similarity 97.6%; Pred. No. 1.6e-82;
Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 APRRLICDSRYLERYLLLEAKENITGCAEHCSELENITVPDTKYNFYAMKRNXXQQA 60
    |||
Db 1 APRRLICDSRYLERYLLLEAKENITGCAEHCSELENITVPDTKYNFYAMKRNXXQQA 60

QY 61 VEWOGATLSEAVLNGQALLVNSQPEPLQHVKAIVGSLTLLPALAOKFAIS 120
    |||
Db 61 VEWOGATLSEAVLNGQALLVNSQPEPLQHVKAIVGSLTLLPALAOKFAIS 120

QY 121 PPDAAAPLRTITADTFKLFYVSNFLGKILKLYTGECRCRIGDR 166
    |||
Db 121 PPDAAAPLRTITADTFKLFYVSNFLGKILKLYTGECRCRIGDR 166

RESULT 2
US-10-014-363-2
; Sequence 2, Application US/10014363
; Patent No. US20020115833A1
; GENERAL INFORMATION:

```

APPLICANT: Burg, Josef
APPLICANT: Engel, Alfred
APPLICANT: Franze, Reinhard
APPLICANT: Hilger, Bernd
APPLICANT: Schurig, Hartmut Ernst
APPLICANT: Tischer, Wilhelm
APPLICANT: Wozny, Manfred
TITLE OF INVENTION: Erythropoietin Conjugates
FILE REFERENCE: Case 20805
CURRENT APPLICATION NUMBER: US/10/014,363
CURRENT FILING DATE: 2001-12-11
NUMBER OF SEQ ID NOS: 5
SOFTWARE: Patentin version 3.1
SEQ ID NO 2
LENGTH: 166
TYPE: PRT
ORGANISM: Homo sapiens
US-10-014-363-2

Query Match 97.9%; Score 825; DB 12; Length 166;
Best Local Similarity 97.6%; Pred. No. 1.6e-82;
Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 APRILICDSRVLELYLLEAKAEENITTCGAHCSINENITVPDITVNFYAMKRNKXGOA 60
DB 1 APRILICDSRVLELYLLEAKAEENITTCGAHCSINENITVPDITVNFYAMKRNKXGOA 60
QY 61 VEVWOGALLSEAVLRGALLVNSSQPWEPIQLHYDKAVSGIRSLTTLRALGAOKEAIS 120
DB 61 VEVWOGALLSEAVLRGALLVNSSQPWEPIQLHYDKAVSGIRSLTTLRALGAOKEAIS 120
QY 121 PPDAASAPLRTITADTFRKLFYVSNFLRGKLYTGEACRTGDR 166
DB 121 PPDAASAPLRTITADTFRKLFYVSNFLRGKLYTGEACRTGDR 166

RESULT 3
US-10-014-363-4
Sequence 4, Application US/10014363
Patent No. US2002011583A1
GENERAL INFORMATION:
APPLICANT: Burg, Josef
APPLICANT: Engel, Alfred
APPLICANT: Franze, Reinhard
APPLICANT: Hilger, Bernd
APPLICANT: Schurig, Hartmut Ernst
APPLICANT: Tischer, Wilhelm
APPLICANT: Wozny, Manfred
TITLE OF INVENTION: Erythropoietin Conjugates
FILE REFERENCE: Case 20805
CURRENT APPLICATION NUMBER: US/10/014,363
CURRENT FILING DATE: 2001-12-11
NUMBER OF SEQ ID NOS: 5
SOFTWARE: Patentin version 3.1
SEQ ID NO 4
LENGTH: 169
TYPE: PRT
ORGANISM: CHO/dhfr-
US-10-014-363-4

Query Match 97.9%; Score 825; DB 12; Length 169;
Best Local Similarity 97.6%; Pred. No. 1.7e-82;
Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 APRILICDSRVLELYLLEAKAEENITTCGAHCSINENITVPDITVNFYAMKRNKXGOA 60
DB 4 APRILICDSRVLELYLLEAKAEENITTCGAHCSINENITVPDITVNFYAMKRNKXGOA 63
QY 61 VEVWOGALLSEAVLRGALLVNSSQPWEPIQLHYDKAVSGIRSLTTLRALGAOKEAIS 120
DB 64 VEVWOGALLSEAVLRGALLVNSSQPWEPIQLHYDKAVSGIRSLTTLRALGAOKEAIS 123
QY 121 PPDAASAPLRTITADTFRKLFYVSNFLRGKLYTGEACRTGDR 166

DB 124 PPDAASAPLRTITADTFRKLFYVSNFLRGKLYTGEACRTGDR 169

RESULT 4
US-10-014-363-3
Sequence 3, Application US/10014363
Patent No. US2002011583A1
GENERAL INFORMATION:
APPLICANT: Burg, Josef
APPLICANT: Engel, Alfred
APPLICANT: Franze, Reinhard
APPLICANT: Hilger, Bernd
APPLICANT: Schurig, Hartmut Ernst
APPLICANT: Tischer, Wilhelm
APPLICANT: Wozny, Manfred
TITLE OF INVENTION: Erythropoietin Conjugates
FILE REFERENCE: Case 20805
CURRENT APPLICATION NUMBER: US/10/014,363
CURRENT FILING DATE: 2001-12-11
NUMBER OF SEQ ID NOS: 5
SOFTWARE: Patentin version 3.1
SEQ ID NO 3
LENGTH: 174
TYPE: PRT
ORGANISM: CHO/dhfr-
US-10-014-363-3

Query Match 97.9%; Score 825; DB 12; Length 174;
Best Local Similarity 97.6%; Pred. No. 1.7e-82;
Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 APRILICDSRVLELYLLEAKAEENITTCGAHCSINENITVPDITVNFYAMKRNKXGOA 60
DB 9 APRILICDSRVLELYLLEAKAEENITTCGAHCSINENITVPDITVNFYAMKRNKXGOA 68
QY 61 VEVWOGALLSEAVLRGALLVNSSQPWEPIQLHYDKAVSGIRSLTTLRALGAOKEAIS 120
DB 69 VEVWOGALLSEAVLRGALLVNSSQPWEPIQLHYDKAVSGIRSLTTLRALGAOKEAIS 128
QY 121 PPDAASAPLRTITADTFRKLFYVSNFLRGKLYTGEACRTGDR 166
DB 129 PPDAASAPLRTITADTFRKLFYVSNFLRGKLYTGEACRTGDR 174

RESULT 5
US-10-014-363-5
Sequence 5, Application US/10014363
Patent No. US2002011583A1
GENERAL INFORMATION:
APPLICANT: Burg, Josef
APPLICANT: Engel, Alfred
APPLICANT: Franze, Reinhard
APPLICANT: Hilger, Bernd
APPLICANT: Schurig, Hartmut Ernst
APPLICANT: Tischer, Wilhelm
APPLICANT: Wozny, Manfred
TITLE OF INVENTION: Erythropoietin Conjugates
FILE REFERENCE: Case 20805
CURRENT APPLICATION NUMBER: US/10/014,363
CURRENT FILING DATE: 2001-12-11
NUMBER OF SEQ ID NOS: 5
SOFTWARE: Patentin version 3.1
SEQ ID NO 5
LENGTH: 174
TYPE: PRT
ORGANISM: CHO/dhfr-
US-10-014-363-5

Query Match 97.9%; Score 825; DB 12; Length 174;
Best Local Similarity 97.6%; Pred. No. 1.7e-82;
Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
QY 1 APPRLICDSRVLERYLLEAKAEENITTCGAHCSLNNITVPDTKYNFYAMKRNKXSQQA 60
db 9 APPRLICDSRVLERYLLEAKAEENITTCGAHCSLNNITVPDTKYNFYAMKRMVEVGOQA 68
QY 61 VEWOGALLSEAVLRGQALLVNSQPWEPLQJHVDKAVSGLSLTLLRALGAQKEAIS 120
db 69 VEWOGALLSEAVLRGQALLVNSQPWEPLQJHVDKAVSGLSLTLLRALGAQKEAIS 128
QY 121 PPDASAPLRTITADTFRKLFRVYSNPLRGKLTLYTGECRTGDR 166
db 129 PPDASAPLRTITADTFRKLFRVYSNPLRGKLTLYTGECRTGDR 174

RESULT 6
US-09-853-731-1
; Sequence 1, Application US/09853731
; Patent No. US20020037841A1
; GENERAL INFORMATION:
; APPLICANT: Papadimitriou, Apollon
; TITLE OF INVENTION: Erythropoietin Composition
; FILE REFERENCE: 20619 US
; CURRENT APPLICATION NUMBER: US/09/853,731
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: EP/00110355.5
; PRIOR FILING DATE: 2000-05-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-853-731-1

Query Match 97.3%; Score 820; DB 10; Length 165;
Best Local Similarity 97.6%; Pred. No. 5,6e-82;
Matches 161; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 APPRLICDSRVLERYLLEAKAEENITTCGAHCSLNNITVPDTKYNFYAMKRNKXSQQA 60
db 1 APPRLICDSRVLERYLLEAKAEENITTCGAHCSLNNITVPDTKYNFYAMKRMVEVGOQA 60
QY 61 VEWOGALLSEAVLRGQALLVNSQPWEPLQJHVDKAVSGLSLTLLRALGAQKEAIS 120
db 61 VEWOGALLSEAVLRGQALLVNSQPWEPLQJHVDKAVSGLSLTLLRALGAQKEAIS 120
QY 121 PPDASAPLRTITADTFRKLFRVYSNPLRGKLTLYTGECRTGDR 165
db 121 PPDASAPLRTITADTFRKLFRVYSNPLRGKLTLYTGECRTGDR 165

RESULT 7
US-10-014-363-1
; Sequence 1, Application US/10014363
; Patent No. US20020115833A1
; GENERAL INFORMATION:
; APPLICANT: Burg, Josef
; APPLICANT: Engel, Alfred
; APPLICANT: Franze, Reinhard
; APPLICANT: Hilger, Bernd
; APPLICANT: Schurig, Hartmut Ernst
; APPLICANT: Tischer, Wilhelm
; APPLICANT: Wozny, Manfred
; TITLE OF INVENTION: Erythropoietin Conjugates
; FILE REFERENCE: Case 20805
; CURRENT APPLICATION NUMBER: US/10/014,363
; CURRENT FILING DATE: 2001-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-014-363-1
```

```
Query Match 97.3%; Score 820; DB 12; Length 165;
Best Local Similarity 97.6%; Pred. No. 5,6e-82;
Matches 161; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 APPRLICDSRVLERYLLEAKAEENITTCGAHCSLNNITVPDTKYNFYAMKRNKXSQQA 60
db 1 APPRLICDSRVLERYLLEAKAEENITTCGAHCSLNNITVPDTKYNFYAMKRMVEVGOQA 60
QY 61 VEWOGALLSEAVLRGQALLVNSQPWEPLQJHVDKAVSGLSLTLLRALGAQKEAIS 120
db 61 VEWOGALLSEAVLRGQALLVNSQPWEPLQJHVDKAVSGLSLTLLRALGAQKEAIS 120
QY 121 PPDASAPLRTITADTFRKLFRVYSNPLRGKLTLYTGECRTGDR 165
db 121 PPDASAPLRTITADTFRKLFRVYSNPLRGKLTLYTGECRTGDR 165

RESULT 8
US-09-864-761-48502
; Sequence 48502, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 48502
; LENGTH: 60
; TYPE: PRT
```

ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AF033356.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.93
OTHER INFORMATION: SWISSPROT HIT: P01588, EVALU 9.00e-28
OTHER INFORMATION: EST_HUMAN HIT: AA662379.1, EVALU 3.00e-10
US-09-864-761-48502

Query Match 33.7%; Score 284; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 2.8e-24;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 58 OQAVVWGLALSLFVAVLGGALLVNSSQWPEPLQHLVDKAVSGLRSLTTLRAAGQ 115
DB 3 OQAVVWGLALSLFVAVLGGALLVNSSQWPEPLQHLVDKAVSGLRSLTTLRAAGQ 60

RESULT 9

US-09-975-063-2
Sequence 2, Application US/09975063
Patent No. US20020045255A1
GENERAL INFORMATION:

APPLICANT: POWELL, Jerry S.
TITLE OF INVENTION: HUMAN ERYTHROPOIETIN GENE: HIGH LEVEL
EXPRESSION IN STABLY TRANSFECTED MAMMALIAN CELLS

NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:

ADDRESSEE: Dorsey & Whitney, LLP
STREET: Suite 3400, 1420 Fifth Avenue, U.S. Bank Centre
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.7/7/8
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/975.063
FILING DATE: 10-Oct-2001
ATTORNEY/AGENT INFORMATION:
NAME: Roberts, Mark W.
REGISTRATION NUMBER: 46,160
REFERENCE/DOCKET NUMBER: 500582.03 (112893.109)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-903-8728
TELEFAX: 206-903-8820
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-975-063-2

Query Match 13.2%; Score 111; DB 10; Length 30;
Best Local Similarity 83.3%; Pred. No. 7e-06;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLERYLLLEAKENITTCGA 30
DB 1 APXRLILDSRVLERYLLLEAKENITTCGA 30

US-10-011-858-2
Sequence 2, Application US/10011858
Patent No. US20020137145A1
GENERAL INFORMATION:

APPLICANT: POWELL, Jerry S.

TITLE OF INVENTION: HUMAN ERYTHROPOIETIN GENE: HIGH LEVEL
EXPRESSION IN STABLY TRANSFECTED MAMMALIAN CELLS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dorsey & Whitney, LLP
STREET: Suite 3400, 1420 Fifth Avenue, U.S. Bank Centre
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS-WORD
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/011.858
FILING DATE: 05-No. US20020137145A1-2001
ATTORNEY/AGENT INFORMATION:
NAME: Roberts, Mark W.
REGISTRATION NUMBER: 46,160
REFERENCE/DOCKET NUMBER: 500582.13
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-903-8728
TELEFAX: 206-903-8820
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-011-858-2

Query Match 13.2%; Score 111; DB 12; Length 30;
Best Local Similarity 83.3%; Pred. No. 7e-06;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLERYLLLEAKENITTCGA 30
DB 1 APXRLILDSRVLERYLLLEAKENITTCGA 30

RESULT 11

US-09-919-703-9
Sequence 9, Application US/09919703
Patent No. US20020165129A1
GENERAL INFORMATION:

APPLICANT: Krystal, Gerald
TITLE OF INVENTION: Peptides and Their Use to Ameliorate
Cell Death
FILE REFERENCE: 50216/003004
CURRENT APPLICATION NUMBER: US/09/919.703
CURRENT FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: US 09/294,457
PRIOR FILING DATE: 1999-04-19
PRIOR APPLICATION NUMBER: US 08/759,599
PRIOR FILING DATE: 1996-12-05
PRIOR APPLICATION NUMBER: US 60/008,233
PRIOR FILING DATE: 1995-12-06
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9

LENGTH: 20
TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Synthetic polypeptide
US-09-919-703-9

Query Match 12.8%; Score 108; DB 9; Length 20;


```

: APPLICANT: Ohide, Akiko
: TITLE OF INVENTION: TUMOR SUPPRESSOR GENE
: FILE REFERENCE: 06501-091001
: CURRENT APPLICATION NUMBER: US/10/045,815
: CURRENT FILING DATE: 2001-10-26
: PRIOR APPLICATION NUMBER: PCT/JP00/02731
: PRIOR FILING DATE: 2000-04-26
: PRIOR APPLICATION NUMBER: JP 11/118806
: PRIOR FILING DATE: 1999-04-26
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 6
: LENGTH: 747
: TYPE: PRT
: ORGANISM: Mus musculus
: US-10-045-815-6

```

```

Query Match      8.2%; Score 69.5; DB 9; Length 747;
Best Local Similarity 25.4%; Pred. No. 20;
Matches 30; Conservative 12; Mismatches 49; Indels 27; Gaps 6;

OY 27 TGCATGCSLNERITVPDTKVNPFYAKRNKXKQAAVEVWGLALTSNAVLRGOALL----- 81
    ||| | : | | | : | | | | | | | | | | |
Db 21 TCGCF-SEPGMDVAPDL---LYA---EGTAIYSRRDMPGVYLNMERALRSRAALLRL 73
    : | | | | | | | | | | | | | | | | |
OY 82 -----VNSSQWEP-LQLHVDKAVS-----GLRSITTLRLALGAQKEAISPPDA 124
    : | | | | | | | | | | | | | | | | |
Db 74 RCRTRCATELPMAPDLIDLGPPSLSDPGAAALHDLRFEGAVLRRACLRCLGPPSA 131
    : | | | | | | | | | | | | | | | | |

```

Search completed: January 8, 2003, 05:17:50
 Job time : 31.6685 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 03:39:48 ; Search time 11.0975 Seconds
(without alignments)
440.118 Million cell updates/sec

Title: US-09-813-775C-18

Perfect score: 843
Sequence: 1 APPRLICDSRVLEERLEAK.....NFLRGKIKLYTGACRTGDR 166

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	827	98.1	166	5	PCT-US94-04361-45
2	825	97.9	166	1	US-08-318-193-70
3	825	97.9	166	4	US-09-604-871-2
4	825	97.9	166	5	PCT-US94-04361-37
5	825	97.9	193	1	US-07-903-220-1
6	825	97.9	193	2	US-08-883-795A-34
7	822	97.5	412	4	US-09-366-009-34
8	820	97.3	165	4	US-09-604-871-1
9	743.5	88.2	165	5	PCT-US94-04361-38
10	741.5	88.0	165	5	PCT-US94-04361-39
11	695	82.4	166	5	PCT-US94-04361-44
12	692	82.1	166	5	PCT-US94-04361-41
13	676	80.2	167	5	PCT-US94-04361-40
14	665.5	78.9	167	5	PCT-US94-04361-42
15	665	78.9	168	5	PCT-US94-04361-43
16	182	21.6	36	5	PCT-US94-04361-50
17	128	15.2	27	5	PCT-US94-04361-51
18	114.5	13.6	34	5	PCT-US94-04361-52
19	108	12.8	20	2	US-08-759-599-9
20	108	12.8	20	4	US-09-294-457-9
21	107	12.7	21	5	PCT-US94-04361-49
22	101	12.0	20	5	PCT-US94-04361-58
23	96	11.4	332	4	US-08-875-533-67
24	95	11.3	312	4	US-08-875-533-66
25	93	11.0	153	4	US-08-875-533-65
26	93	11.0	165	1	US-08-413-803-29
27	93	11.0	174	3	US-08-471-045-56

28	93	11.0	174	3	US-08-469-712A-56	Sequence 56, Appl
29	93	11.0	174	4	US-08-446-871-56	Sequence 56, Appl
30	93	11.0	174	4	US-08-468-910-56	Sequence 56, Appl
31	93	11.0	174	4	US-08-761-907-56	Sequence 56, Appl
32	93	11.0	174	5	PCT-US95-03776-27	Sequence 27, Appl
33	93	11.0	195	1	US-08-388-779A-4	Sequence 4, Appl
34	93	11.0	195	1	US-08-591-070A-4	Sequence 4, Appl
35	93	11.0	195	2	US-08-927-855-4	Sequence 4, Appl
36	93	11.0	285	4	US-08-875-533-69	Sequence 69, Appl
37	93	11.0	285	4	US-08-875-533-70	Sequence 70, Appl
38	93	11.0	288	4	US-08-875-533-71	Sequence 71, Appl
39	93	11.0	288	4	US-08-875-533-72	Sequence 73, Appl
40	93	11.0	332	4	US-08-875-533-72	Sequence 22, Appl
41	93	11.0	332	5	PCT-US95-03776-25	Sequence 25, Appl
42	93	11.0	353	1	US-08-330-517-2	Sequence 2, Appl
43	93	11.0	353	1	US-08-347-029-4	Sequence 4, Appl
44	93	11.0	353	1	US-08-380-779A-2	Sequence 2, Appl
45	93	11.0	353	1	US-08-484-246-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
PCT-US94-04361-45
Sequence 45, Application PC/TUS9404361
GENERAL INFORMATION:
APPLICANT: Brigham and Women's Hospital
APPLICANT: 75 Francis Street
APPLICANT: Boston, MA 02115
APPLICANT: Bunn, H. Franklin
APPLICANT: Wen, Danyl
TITLE OF INVENTION: Erythropoietin Mutains With Enhanced
TITLE OF INVENTION: Activity
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04361
FILING DATE: Herewith
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/049,802
FILING DATE: 21-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Cimballa, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0627.336PC01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acids
TYPE: amino acid
TOPOLOGY: both
PCT-US94-04361-45

Query Match 98.1% Score 827; DB 5; Length 166;
Best Local Similarity 97.6% Pred. No. 1e-96;
Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy	1	APPRICDSRYLEERYLLLEAKFAENITTCGACHCGLNINITYPDTKVFYAMKRNXSQQA	60
Db	1	APPRICDSRYLEERYLLLEAKFAENITTCGACHCGLNINITYPDTKVFYAMKRNREGQQA	60
Qy	61	VEVWOGALLSEAVLRQALLVNSQSOMPELTQIHYDKAVSGLRSTLTLLRALGQRPAS	120
Db	61	VEVWOGALLSEAVLRQALLVNSQSOMPELTQIHYDKAVSGLRSTLTLLRALGQRPAS	120
Qy	121	PPDAASAAPLRTITADYDFRKLTFRYYSNPLRGKLTLYTGEACRIGDR	166
Db	121	PPDAASAAPLRTITADYDFRKLTFRYYSNPLRGKLTLYTGEACRIGDR	166

RESULT 2

```

US-08-318-193-70
: Sequence 70, Application US/08318193
: Patent No. 5641663
: GENERAL INFORMATION:
: APPLICANT: GARVIN, Robert T.
: APPLICANT: MALEK, Lawrence T.
: TITLE OF INVENTION: AN EXPRESSION SYSTEM FOR THE SECRETION
: TITLE OF INVENTION: OF BIOACTIVE HUMAN GRANULOCYTE MACROPHAGE COLONY
: TITLE OF INVENTION: STIMULATING FACTOR (GM-CSF) AND OTHER HETEROLOGOUS
: TITLE OF INVENTION: PROTEINS FROM STREPTOMYCES
: NUMBER OF SEQUENCES: 91
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 1800 Diagonal Road, Suite 500
: CITY: Alexandria
: STATE: Virginia
: COUNTRY: USA
: ZIP: 22313-0299
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/318,193
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/935,314
: FILING DATE:
: APPLICATION NUMBER: US 07/224,568
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 18740/116 CACO
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703)836-9300
: TELEFAX: (703)685-4109
: TELEX: 899149
: INFORMATION FOR SEQ ID NO: 70:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 166 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-318-193-70

```

Query Match	97.9%	Score 825;	DB 1;	Length 166;
Best Local Similarity	97.6%	Pred. No. 1.8e-96;		
Matches 162; Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;

Qy	Db	Qy	Db
1	1	61	61
APPLICDSRYLERYLLEAKAEENITTCGAHCSSLNINITYPDTKYVFAMKRNKXSQQA	APPLICDSRYLERYLLEAKAEENITTCGAHCSSLNINITYPDTKYVFAMKRNKREYQQA	VEVWOGALLSEANLRQCALLVNSSQWPEPLQLHYDKAVSGLSLITLLPALCAQKEAIS	VEVWOGALLSEANLRQCALLVNSSQWPEPLQLHYDKAVSGLSLITLLPALCAQKEAIS
120	60	120	120

QY 121 PPDASAPLRTITADTFKRLFRVYSNFLRGKLLKLTGEACRPGDR 166
 |||||
 Db 121 PPDASAPLRTITADTFKRLFRVYSNFLRGKLLKLTGEACRPGDR 166

RESULT 3
US-09-604-871-2

```

? Sequence 2, Application US/09b04871
? Patent No. 6340742
? GENERAL INFORMATION:
? APPLICANT: Burd, Josef
? APPLICANT: Hilger, Bernd
? APPLICANT: Joesel, Hans-Peter
? TITLE OF INVENTION: EYTHHPOLEIETIN CONJUGATES
? FILE REFERENCE: 1098 nonprovisional
? CURRENT APPLICATION NUMBER: US/09/604,871
? PRIOR FILING DATE: 2000-06-28
? PRIOR FILING DATE: 60/151,454
? PRIOR FILING DATE: 1999-08-30
? PRIOR APPLICATION NUMBER: 60/147,452
? PRIOR FILING DATE: 1999-08-05
? PRIOR APPLICATION NUMBER: 60/142,243
? PRIOR FILING DATE: 1999-07-02
? NUMBER OF SEQ ID NOS: 3
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 2
?
? LENGTH: 166
? TYPE: PRT
? ORGANISM: Homo sapiens
? OS-09-604-871-2

```

Query Match	97.9%	Score 825; DB 4;	Length 166;
Best Local Similarity	97.6%	Pred. No. 1.8e-96;	
Matches 162; Conservative	0;	Mismatches 4;	Indels 0; Gaps 0;

QY	1	APPRLIDSRILEXYLLLEAKAEENITTCGAHCGLNENITYPDTKVFVAMKRKXSKQA	60
Db	1	APPLRIDSRILRYELLLEAKEAENITTCGAHCSLNENITYPDTRKNFVAMKRMEVQQA	60
QY	61	VEWVGALLSEAVLRGCAQLLVNSQSOPWEPIQLHYDKVASGRSLTTLRALAQKAIS	120
Db	61	VEWVGALLSEAVLRGCAQLLVNSQSOPWEPIQLHYDKVASGRSLTTLRLRAGQKAIS	120
QY	121	PDDAASAPLRTITADTFRKLFERYYSNFLRGKLKLYTGCAQCRTDGR	166
Db	121	PPDAASAPLRTITADTFRKLFERYYSNFLRGKLKLYTGCAQCRTGDR	166

RESULT 4
PCT-US94-04361-37

```

? Sequence 37, Application PC/TUS9404361
?
? GENERAL INFORMATION:
?
? APPLICANT: Brigham and Women's Hospital
?
? APPLICANT: 75 Francis Street
?
? APPLICANT: Boston, MA 02115
?
? APPLICANT: Bunn, H. Franklin
?
? APPLICANT: Wen, Danyi
?
? APPLICANT: Showers, Mark O.
?
? TITLE OF INVENTION: Erythropoietin Mucins With Enhanced
?
? TITLE OF INVENTION: Activity
?
? NUMBER OF SEQUENCES: 59

```

```

1  CURRENT APPLICATION DATA:
2  APPLICATION NUMBER: PCT/US94/04361
3  FILING DATE: Herewith
4  CLASSIFICATION:
5  PRIOR APPLICATION DATA:
6  APPLICATION NUMBER: 08/049,802
7  FILING DATE: 21-APR-1993
8  ATTORNEY/AGENT INFORMATION:
9  NAME: Cimdala, Michele A.
10 REGISTRATION NUMBER: 33,851
11 REFERENCE/DOCKET NUMBER: 0627, 336PC01
12 TELECOMMUNICATION INFORMATION:
13 TELEPHONE: (402) 371-2600
14 TELEFAX: (402) 371-2540
15 INFORMATION FOR SEQ ID NO: 37:
16 SEQUENCE CHARACTERISTICS:
17     LENGTH: 166 amino acids
18     TYPE: amino acid
19     TOPOLOGY: both
20 PCT-US94-04361-37

```

Query Match	97.98	Score 825	DB 5	Length 166
Best Local Similarity	97.68	Pred. No.	1.8e-96	
Matches 162	Conservative	0	Mismatches 4	Indels 0
				Gaps 0

QY	1	APPRLCD\$VLE\$YLLLEAK\$EAE\$ITTG\$CA\$H\$G\$I\$N\$ENIT\$V\$P\$D\$T\$K\$V\$F\$Y\$A\$M\$K\$R\$X\$S\$X\$O\$A	60
Db	1	APPRICDS\$VLE\$YLLLEAK\$EAE\$NITTG\$CA\$H\$G\$I\$N\$ENIT\$V\$P\$D\$T\$K\$V\$F\$Y\$A\$M\$K\$R\$E\$V\$G\$O\$A	60
QY	61	VEWNOG\$LL\$SEAV\$LRG\$ALL\$VNS\$Q\$P\$E\$P\$LO\$LV\$H\$D\$K\$A\$V\$G\$R\$S\$T\$LL\$P\$AL\$C\$A\$O\$K\$R\$A\$S	120
Db	61	VEWNOG\$LL\$SEAV\$LRG\$ALL\$VNS\$Q\$P\$E\$P\$LO\$LV\$H\$D\$K\$A\$V\$G\$R\$S\$T\$LL\$P\$AL\$C\$A\$O\$K\$R\$A\$S	120
QY	121	PPDA\$A\$A\$P\$E\$R\$T\$IT\$AD\$T\$F\$R\$K\$L\$F\$R\$Y\$S\$N\$F\$L\$G\$R\$K\$L\$K\$Y\$T\$G\$E\$A\$C\$R\$T\$G\$D\$R	166
Db	121	PPDA\$A\$A\$P\$E\$R\$T\$IT\$AD\$T\$F\$R\$K\$L\$F\$R\$Y\$S\$N\$F\$L\$G\$R\$K\$L\$K\$Y\$T\$G\$E\$A\$C\$R\$T\$G\$D\$R	166

RESULT 5
 US-07-903-220-1
 : Sequence 1, Application US/07903220
 Patent No. 532837
 : GENERAL INFORMATION:
 APPLICANT: Hewick, Rodney M.
 TITLE OF INVENTION: METHOD FOR THE PURIFICATION OF
 TITLE OF INVENTION: ERYTHROPOIETIN AND ERYTHROPOIETIN COMPOSITION
 NUMBER OF SEQUENCES: 1
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Paul H. Heller
 STREET: Kenyon & Kenyon, One Broadway
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/903,220
 FILING DATE: 19920731
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Brown, Scott A.
 REGISTRATION NUMBER: 32,724
 REFERENCE/DOCKET NUMBER: 1248/27
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 429-1776
 TELEFAX: (202) 429-0796
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 193 amino acids

```

; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-07-903-220-1

```

Query Match	97.9%	Score 825	DB 1	Length 193
Best Local Similarity	97.6%	Pred. No. 2.3e-96		
Matches 162	Conservative 0	Mismatches 4	Indels 0	Gaps 0

[illegible][illegible]

```

1      RESULT 6
2      US-08-883-795A-34
3      : Sequence 34, Application US/08883795A
4      : Patent No. 5985607
5      : GENERAL INFORMATION:
6      : APPLICANT: Delcuve, Genevieve
7      : APPLICANT: Awang, Gregor
8      : TITLE OF INVENTION: Recombinant DNA Molecules and Expression
9      : TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
10     : NUMBER OF SEQUENCES: 39
11     : CORRESPONDENCE ADDRESS:
12     : ADDRESSEE: BERSKIN & PARR
13     : STREET: 40 King Street West
14     : CITY: Toronto
15     : STATE: Ontario
16     : COUNTRY: Canada
17     : ZIP: M5H 3Y2
18     :
19     : COMPUTER READABLE FORM:
20     : MEDIUM TYPE: Floppy disk
21     : COMPUTER: IBM PC compatible
22     : OPERATING SYSTEM: PC-DOS/MS-DOS
23     : SOFTWARE: PatentIn Release #1.0, Version #1.25
24     : CURRENT APPLICATION DATA:
25     : APPLICATION NUMBER: US/08/883,795A
26     : FILING DATE: 27-JUN-1997
27     : CLASSIFICATION: 435
28     : ATTORNEY/AGENT INFORMATION:
29     : NAME: Gravelle, Micheline
30     : REGISTRATION NUMBER: 40,261
31     : REFERENCE/DOCKET NUMBER: 7841-062
32     : TELECOMMUNICATION INFORMATION:
33     : TELEPHONE: (416) 364-7311
34     : TELEFAX: (416) 361-1398
35     : INFORMATION FOR SEQ ID NO: 34:
36     : SEQUENCE CHARACTERISTICS:
37     : LENGTH: 193 amino acids
38     : TYPE: amino acid
39     : TOPOLOGY: linear
40     :
41     : MOLECULE TYPE: protein
42     :
43     : US-08-883-795A-34

```

Query Match	97.98;	Score 825;	DB 2;	Length 193;
Best Local Similarity	97.68;	Pred. No. 2.3e-96;		
Matches 162;	Conservative	0;	Mismatches 4;	Indels 0;
				Gaps 0;

Qy 1 APRRLICDSRVLEERYLLFAKAEANTTGCACHSCLNENITVPPTKVNFYAMKRNNXQQQA 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 28 APRRLICDSRVLEERYLLFAKAEANTTGCACHSCLNENITVPPTKVNFYAMKRMEVGQQA 87


```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04361
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/049,802
FILING DATE: 21-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Cimbalia, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0627.336PC01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 165 amino acids
TYPE: amino acid
TOPOLOGY: both
PCT-US94-04361-38

Query Match      88.2%; Score 743.5; DB 5; Length 165;
Best Local Similarity 89.2%; Pred. No. 3.8e-86;
Matches 148; Conservative 7; Mismatches 10; Indels 1; Gaps 1;

OY 1 APRRLICDSRVLEERYLLLEAKAEENITTCGAHCSLNENITVPDTKVFYAMKRNXSXQA 60
    |||
DB 1 APRRLICDSRVLEERYLLLEAKAEENITTCGAHCSLNENITVPDTKVFYAMKRNXSXQA 60
    |||
OY 61 VEWOGIALLEAVLRGOALLVNSSQPWEPLQLHVDKAVGSLSTLLRALGAKRAIS 120
    |||
DB 61 VEWOGIALLEAVLRGOALLVNSSQPWEPLQLHMDKAIISGLSITLLRALGAQ-BAIS 119
    |||
OY 121 PPDAASAPLRTITADTFKRLFRVYSNPLRGKLYTGECACRGDR 166
    |||
DB 120 LPDAASAPLRTITADTFKRLFRVYSNPLRGKLYTGECACRGDR 165
    |||

RESULT 10
PCT-US94-04361-39
Sequence 39, Application PC/TUS9404361
GENERAL INFORMATION:
APPLICANT: Brigham and Women's Hospital
APPLICANT: 75 Francis Street
APPLICANT: Boston, MA 02115
APPLICANT: Bunn, H. Franklin
APPLICANT: Men, Danyl
APPLICANT: Showers, Mark O.
TITLE OF INVENTION: Erythropoietin Mutelins With Enhanced
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04361
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/049,802
FILING DATE: 21-APR-1993
```

```
ATTORNEY/AGENT INFORMATION:
NAME: Cimbalia, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0627.336PC01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 165 amino acids
TYPE: amino acid
TOPOLOGY: both
PCT-US94-04361-39

Query Match      88.0%; Score 741.5; DB 5; Length 165;
Best Local Similarity 88.6%; Pred. No. 6.7e-86;
Matches 147; Conservative 8; Mismatches 10; Indels 1; Gaps 1;

OY 1 APRRLICDSRVLEERYLLLEAKAEENITTCGAHCSLNENITVPDTKVFYAMKRNXSXQA 60
    |||
DB 1 APRRLICDSRVLEERYLLLEAKAEENITTCGAHCSLNENITVPDTKVFYAMKRNXSXQA 60
    |||
OY 61 VEWOGIALLEAVLRGOALLVNSSQPWEPLQLHVDKAVGSLSTLLRALGAKRAIS 120
    |||
DB 61 VEWOGIALLEAVLRGOALLVNSSQPWEPLQLHMDKAIISGLSITLLRALGAQ-BAIS 119
    |||
OY 121 PPDAASAPLRTITADTFKRLFRVYSNPLRGKLYTGECACRGDR 166
    |||
DB 120 LPDAASAPLRTITADTFKRLFRVYSNPLRGKLYTGECACRGDR 165
    |||

RESULT 11
PCT-US94-04361-44
Sequence 44, Application PC/TUS9404361
GENERAL INFORMATION:
APPLICANT: Brigham and Women's Hospital
APPLICANT: 75 Francis Street
APPLICANT: Boston, MA 02115
APPLICANT: Bunn, H. Franklin
APPLICANT: Men, Danyl
APPLICANT: Showers, Mark O.
TITLE OF INVENTION: Erythropoietin Mutelins With Enhanced
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04361
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/049,802
FILING DATE: 21-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Cimbalia, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0627.336PC01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acids
```

TYPE: amino acid
TOPOLOGY: both
PCT-US94-04361-44

Query Match
Best Local Similarity 82.4%; Score 695; DB 5; Length 166;
Matches 137; Conservative 8; Mismatches 21; Indels 0; Gaps 0;

QY 1 APPRLICDSRVLEERYLLAKEAENITTCGAHCSENENITVPDITVNFYAMKRNKXSOQA 60
DB 1 APPRLICDSRVLEERYLLAKEAENITTCGAHCSENENITVPDITVNFYAMKRNKXSOQA 60

QY 61 VEWOGGLALLSEAVLRGALLVNSQPEPQLQHYDKAVSGLSRSLTTLRALGAQKEAIS 120
DB 61 VEWOGGLALLSEAVLRGALLVNSQPEPQLQHYDKAVSGLSRSLTTLRALGAQKEAIS 120

QY 121 PPDASAPLRITITADTFPRKLFRRVSNFLRGKLTLYTGECRGTGR 166
DB 121 PPDASAPLRITITADTFPRKLFRRVSNFLRGKLTLYTGECRGTGR 166

RESULT 12

PCT-US94-04361-41
Sequence 41, Application PC/TUS9404361
GENERAL INFORMATION:
APPLICANT: Brigham and Women's Hospital
APPLICANT: 75 Francis Street
APPLICANT: Boston, MA 02115
APPLICANT: Bunn, H. Franklin
APPLICANT: Men, Danyl
TITLE OF INVENTION: Erythropoietin Mutelins With Enhanced
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04361
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/049,802
FILING DATE: 21-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Cimbal, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0627,336PC01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acids
TYPE: amino acid
TOPOLOGY: both
PCT-US94-04361-41

Query Match
Best Local Similarity 82.1%; Score 692; DB 5; Length 166;
Matches 135; Conservative 12; Mismatches 19; Indels 0; Gaps 0;

QY 1 APPRLICDSRVLEERYLLAKEAENITTCGAHCSENENITVPDITVNFYAMKRNKXSOQA 60
DB 1 APPRLICDSRVLEERYLLAKEAENITTCGAHCSENENITVPDITVNFYAMKRNKXSOQA 60

DB 1 APPRLICDSRVLEERYLLAKEAENITTCGAHCSENENITVPDITVNFYAMKRNKXSOQA 60
QY 61 VEWOGGLALLSEAVLRGALLVNSQPEPQLQHYDKAVSGLSRSLTTLRALGAQKEAIS 120
DB 61 VEWOGGLALLSEAVLRGALLVNSQPEPQLQHYDKAVSGLSRSLTTLRALGAQKEAIS 120

QY 121 PPDASAPLRITITADTFPRKLFRRVSNFLRGKLTLYTGECRGTGR 166
DB 121 PPDASAPLRITITADTFPRKLFRRVSNFLRGKLTLYTGECRGTGR 166

RESULT 13

PCT-US94-04361-40
Sequence 40, Application PC/TUS9404361
GENERAL INFORMATION:
APPLICANT: Brigham and Women's Hospital
APPLICANT: 75 Francis Street
APPLICANT: Boston, MA 02115
APPLICANT: Bunn, H. Franklin
APPLICANT: Men, Danyl
TITLE OF INVENTION: Erythropoietin Mutelins With Enhanced
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04361
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/049,802
FILING DATE: 21-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Cimbal, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0627,336PC01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acids
TYPE: amino acid
TOPOLOGY: both
PCT-US94-04361-40

Query Match
Best Local Similarity 80.2%; Score 676; DB 5; Length 166;
Matches 130; Conservative 14; Mismatches 22; Indels 0; Gaps 0;

QY 1 APPRLICDSRVLEERYLLAKEAENITTCGAHCSENENITVPDITVNFYAMKRNKXSOQA 60
DB 1 APPRLICDSRVLEERYLLAKEAENITTCGAHCSENENITVPDITVNFYAMKRNKXSOQA 60

QY 61 VEWOGGLALLSEAVLRGALLVNSQPEPQLQHYDKAVSGLSRSLTTLRALGAQKEAIS 120
DB 61 VEWOGGLALLSEAVLRGALLVNSQPEPQLQHYDKAVSGLSRSLTTLRALGAQKEAIS 120

QY 121 PPDASAPLRITITADTFPRKLFRRVSNFLRGKLTLYTGECRGTGR 166
DB 121 PPDASAPLRITITADTFPRKLFRRVSNFLRGKLTLYTGECRGTGR 166


```

RESULT 14
PCT-US94-04361-42
; Sequence 42, Application PC/TUS9404361
; GENERAL INFORMATION:
; APPLICANT: Brigham and Women's Hospital
; APPLICANT: 75 Francis Street
; APPLICANT: Boston, MA 02115
; APPLICANT: Bunn, H. Franklin
; APPLICANT: Men, Danyi
; APPLICANT: Showers, Mark O.
; TITLE OF INVENTION: Erythropoietin Mutelins With Enhanced
; TITLE OF INVENTION: Actively
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentln Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04361
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/049, 802
; FILING DATE: 21-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimdala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 0627.336PC01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 167 amino acids
; TYPE: amino acid
; TOPOLOGY: both
;
PCT-US94-04361-42

Query Match 78.9%, Score 665.5; DB 5; Length 167;
Best Local Similarity 79.6%; Pred. No. 2.9e-76;
Matches 133; Conservative 9; Mismatches 24; Indels 1; Gaps 1

QY 1 APPRICDSRVLEERLLLEKAEENTTTCAGHCISLNNENTTVDTKVNFAMKRNXSXQA 60
Db 1 APPRICDSRVLEERLLLEKAEENTTTCAGHCISLNNENTTVDTKVNFAMKRMVEYQQA 60
QY 61 VEWOGALLSEAVLRGQALLVNSQSPWPELDLHVDKAVSGLSRLTTLRALGAQKEATS 120
Db 61 LEVMOGALLSEALFRGQALPANASQPCBALRLHVDKAVSGLSRLTLRALGAQKEALP 120
QY 121 PPDA-SAAPLRTTADTFRKLFVRYSNFLRGKLLLYTGEACRTGDR 166
Db 121 LPDATPSAAPLRTITVDALSKLFRYSNFLRGKTLTYTGEACRGDR 167

RESULT 15
PCT-US94-04361-43
; Sequence 43, Application PC/TUS9404361
; GENERAL INFORMATION:
; APPLICANT: Brigham and Women's Hospital
; APPLICANT: 75 Francis Street
; APPLICANT: Boston, MA 02115
; APPLICANT: Bunn, H. Franklin
; APPLICANT: Wen, Danyi

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 02:01:43 ; Search time 47.6267 Seconds
(without alignments)
718.164 Million cell updates/sec

Title: US-09-813-775c-18

Perfect score: 843
Sequence: 1 APRRLICDSRVLEERYLLEAK.....NFLRGKLIKLYTGSEACTGDR 166

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhch:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP rivirus:*
16: SP bacteriap:*
17: SP archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	675.5	80.1	195	6	O9GKA3	O9GKA3 oryctolagus
2	675.5	80.1	195	6	O9GKA2	O9GKA2 oryctolagus
3	665	78.9	194	6	O9MYM8	O9MYM8 sus scrofa
4	188	22.3	50	11	O9QV40	O9QV40 ratcus sp.
5	87.5	10.4	346	16	O82K24	O82K24 salmoneila
6	87.5	10.4	346	16	O82K25	O82K25 salmoneila
7	84	10.0	339	16	O9HZM7	O9HZM7 pseudomonas
8	82.5	9.8	389	2	O9AN18	O9AN18 bradyrhizob
9	81	9.6	323	16	O8ZDC8	O8ZDC8 yersinia pe
10	80.5	9.5	871	5	O9NB04	O9NB04 drosophila
11	79	9.4	3722	2	P94873	P94873 lysobacter
12	78.5	9.3	383	16	O92QV7	O92QV7 rhizobium m
13	78	9.3	378	17	O82ZK3	O82ZK3 pyrobaculum
14	77.5	9.2	242	16	O8YV76	O8YV76 anabaena sp
15	77.5	9.2	567	10	O9SVL5	O9SVL5 arabidopsis
16	77	9.1	348	16	O86751	O86751 streptomyces

17	77	9.1	451	16	P74054	P74054 synechocyst
18	77	9.1	480	10	O43380	O43380 avena sativ
19	76.5	9.1	379	2	O919F5	O919F5 streptomyces
20	76.5	9.1	1638	2	O87001	O87001 pseudomonas
21	76.5	9.1	2472	16	O91696	O91696 pseudomonas
22	76	9.0	554	2	O9RPH5	O9RPH5 mycobacteri
23	75.5	9.0	451	16	O8XSE8	O8XSE8 ralsionia s
24	74.5	8.8	455	16	O8UBQ9	O8UBQ9 agrobacteri
25	74.5	8.8	756	12	O90155	O90155 trichomonas
26	74.5	8.8	2556	11	O91XW2	O91XW2 mus musculu
27	74	8.8	397	8	O9TNG5	O9TNG5 coriaria te
28	74	8.8	426	3	O9HCD5	O9HCD5 absidia coe
29	74	8.8	815	10	O9PK91	O9PK91 arabidopsis
30	74	8.8	904	5	O9VEX7	O9VEX7 drosophila
31	74	8.8	912	5	O9NHC1	O9NHC1 drosophila
32	74	8.8	912	5	O9NHB9	O9NHB9 drosophila
33	74	8.8	916	5	O9NHC2	O9NHC2 drosophila
34	74	8.8	916	5	O9NHC0	O9NHC0 drosophila
35	74	8.8	1003	10	O9FIR4	O9FIR4 arabidopsis
36	73	8.7	353	16	O8ZGV8	O8ZGV8 yersinia pe
37	73	8.7	454	11	O9D7X0	O9D7X0 mus musculu
38	73	8.7	1481	5	O9NEX0	O9NEX0 caenorhabdi
39	72.5	8.6	637	16	O55413	O55413 synechocyst
40	72.5	8.6	762	16	O9A4Q7	O9A4Q7 caulobacter
41	72.5	8.6	917	2	O9F8D7	O9F8D7 pseudomonas
42	72.5	8.6	917	2	O51722	O51722 pseudomonas
43	72	8.5	318	16	O8XX12	O8XX12 ralsionia s
44	72	8.5	475	4	O8WUT4	O8WUT4 homo sapien
45	72	8.5	896	2	O9AN79	O9AN79 bradyrhizob

ALIGNMENTS

RESULT 1

ID	O9GKA3	PRELIMINARY;	PRT;	195 AA.
AC	O9GKA3;			
DT	01-MAR-2001 (TREMBLrel. 16, Created)			
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Erythropoietin.			
OS	Oryctolagus cuniculus (Rabbit).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.			
OX	NCBI_TaxID=9986;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21290682; PubMed=11396976;			
RA	Vallita A., Wu D., Margalith M., Hobart P.;			
RT	"Rabbit Epo Gene and cDNA: Expression of Rabbit Epo after			
RT	Intramuscular Injection of pDNA."			
RL	Biochem. Biophys. Res. Commun. 284:823-827(2001).			
DR	EMBL; AF290943; AAG36961.1; -.			
DR	HSSP; P01588; ICN4.			
DR	InterPro; IPR001323; EPO_TPO.			
DR	InterPro; IPR003013; Erythropo.			
DR	Pfam; PF00758; EPO_TPO; 1.			
DR	PRINTS; PR00272; ERYTHROPTN.			
DR	PROSITE; PS00817; EPO_TPO; 1.			
SO	SEQUENCE 195 AA; 21053 MW; 0999DA7D852713F3 CRC64;			

Query Match 80.1%; Score 675.5; DB 6; Length 195;
Best Local Similarity 80.2%; Pred. No. 2,3e-60;
Matches 134; Conservative 14; Mismatches 18; Indels 1; Gaps 1;

OY	1	APRRLICDSRVLEERYLLEAKENITTCAGHCISLSENTIVPDTKVNFAAMKRNKXQQA	60
OY	1		60
DB	29	APRRLICDSRVLEERYLLEAKENITTCAGHCISLSENTIVPDTKVNFAAMKRNKXQQA	88
OY	61	VEWOGALILEAVLNGOALLVSSQPMWPELOLHVQKAVGSLSTTLRALGAKENAS	120
OY	1		120
DB	89	VEWOGALILEAVLNGOALLVSSQPMWPELOLHVQKAVGSLSTTLRALGAKENAS	148
DB	1		148

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-21534948 / SCSC1412 / ATCC 700720;
 RX MEDLINE-21534948; PubMed-11677609;
 RA McLelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Potwilk S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 LT2.";
 RL Nature 413:852-856(2001).
 DR EMBL: AE008878; AL22684.1;
 DR InterPro: IPR001761; PeriplabP/LacI.
 DR Pfam: PF00532; Peripla_BP_Like; 1.
 KM Complete proteome.
 SQ SEQUENCE 346 AA; 38440 MW; E37CAB58E49FD716 CRC64;

Query Match 10.4%; Score 87.5; DB 16; Length 346;
 Best Local Similarity 26.7%; Pred. No. 1.3;
 Matches 44; Conservative 22; Mismatches 48; Indels 51; Gaps 9;

OY 10 RVLERYLLAEAEENITTG--CAEHCSLNE--NITVPDTKYNFYAMKRNKXSXQAAVEYQW 65
 DB 217 RNLQEMLEERHPDANVAGSAIAEAAMGEGRNLTPTLIVSEYL-----THQYR 267
 OY 66 GLALSEAVLRGQALLVNSO--PWEPIQLHVDKAVSGIRSLTTLRALGAO--KEAISP 122
 DB 268 GLK-----RGHILMALSDQMAMQ-----GELATQISKIVLQGGPVPENISPP 309
 OY 123 -----DAASAPLRTITADTFKRLFRVYSNFLRGKILTYGEA 160
 DB 310 VLIITHNNADSARVRRLSPGFRPVY-----LYQYTSSEA 344

RESULT 6

ID 0822M5 PRELIMINARY; PRT; 346 AA.
 AC 0822M5;

DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Solute binding receptor protein.
 GN STY3952.

OS *Salmonella typhi*.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Salmonella.

OX NCBI_TaxID=601;

RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-C018;

RX MEDLINE-21534947; PubMed-11677608;

RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,

RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahina M.,

RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,

RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,

RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jorgels K.,

RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,

RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,

RA Whitehead S., Barrett B.G.;

RT "Complete genome sequence of a multiple drug resistant *Salmonella*

LT enterica serovar Typhi CT18.";

RL Nature 413:848-852(2001).

DR EMBL: AL627280; CAD03169.1;

DR InterPro: IPR001761; PeriplabP/LacI.

DR Pfam: PF00532; Peripla_BP_Like; 1.

KM Receptor; Complete proteome.

SQ SEQUENCE 346 AA; 38546 MW; F80FBI688BC3ABF CRC64;

Query Match 10.4%; Score 87.5; DB 16; Length 346;

Best Local Similarity 26.7%; Pred. No. 1.3;
 Matches 44; Conservative 22; Mismatches 48; Indels 51; Gaps 9;

OY 10 RVLERYLLAEAEENITTG--CAEHCSLNE--NITVPDTKYNFYAMKRNKXSXQAAVEYQW 65
 DB 217 RNLQEMLEERHPDANVAGSAIAEAAMGEGRNLTPTLIVSEYL-----THQYR 267
 OY 66 GLALSEAVLRGQALLVNSO--PWEPIQLHVDKAVSGIRSLTTLRALGAO--KEAISP 122
 DB 268 GLK-----RGHILMALSDQMAMQ-----GELATQISKIVLQGGPVPENISPP 309
 OY 123 -----DAASAPLRTITADTFKRLFRVYSNFLRGKILTYGEA 160
 DB 310 VLIITHNNADSARVRRLSPGFRPVY-----LYQYTSSEA 344

RESULT 7

ID 09H2M7 PRELIMINARY; PRT; 339 AA.
 AC 09H2M7;

DT 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE UDP-N-acetylglucosamine 2-aminoglycosamine reductase.

GN MURR OR PA2977.

OS *Pseudomonas aeruginosa*.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

OX NCBI_TaxID=287;

RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 15692 / PA01;

RX MEDLINE=20437337; PubMed=10984043;

RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warriner P.,

RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,

RA Gardner R.L., Goltz L., Tolentino E., Westbrock-Wadman S., Yuan Y.,

RA Brody L.R., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,

RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,

RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;

RT "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an

opportunistic pathogen.";

RL Nature 406:959-964(2000).

DR EMBL: AE004723; AGO06365.1;

DR HSSP: P08373; ZMBR.

DR InterPro: IPR003170; MurB.

DR InterPro: IPR001575; Oxid_FAD_bind.

DR InterPro: IPR000531; TonB_boxC.

DR Pfam: PF01565; FAD_binding_4; 1.

DR Pfam: PF02873; MurB_C; 1.

DR PROSITE: PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.

KM Complete proteome.

SQ SEQUENCE 339 AA; 37627 MW; C0C8EF9F2938FE27 CRC64;

Query Match 10.0%; Score 84; DB 16; Length 339;
 Best Local Similarity 22.4%; Pred. No. 2.9;

Matches 41; Conservative 24; Mismatches 48; Indels 70; Gaps 8;

OY 31 EHCSLNE--NITVPDTKYNFYAMKRNKXSXQAAVEYQWGLALSEAVLRGQALLV----- 82
 DB 7 EHCSLKPYNFTGIDVRLARLAHARDEA-----DVREALALARE--RGLPLVLVGGSNL 58
 OY 83 -----NSQEPER--LQHLVDKAVSGIRSLTTL 109
 DB 59 LLTRDVEALVRMASQGRRIYSDAASVLYEAGEAMDPRVQSLRGLAGLENTLSIT 117
 OY 110 RALGAKAEISPPDAASAPLRTITTA-----DFFKRLFRVYSNFLRGKILTYGCACT 163
 DB 118 -----PCTVGAAPWQONIGAYGVELKDVFDLSITLAL--DROGTILREFDQACRF 163
 OY 164 GDR 166
 DB 164 GYR 166

```

RESULT 8
ID 09AN18 PRELIMINARY; PRT; 389 AA.
AC 09AN18:
DT 01-JUN-2001 (TREMBLREL. 17, Created)
DT 01-JUN-2001 (TREMBLREL. 17, Last sequence update)
DE ID626.
GN ID626.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bradyrhizobium group; Bradyrhizobium.
OX NCBI_TaxID=375;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=110SPC4;
RX MEDLINE=21101824; PubMed=1157954;
RA Hennecke H.;
RT "Potential symbiosis-specific genes uncovered by sequencing a 410-kb
RT DNA region of the Bradyrhizobium japonicum chromosome.";
RL J. Bacteriol. 183:1405-1412(2001).
DR EMBL: AF322013; AAG60960.1; -
SQ SEQUENCE 389 AA; 42316 MW; 6E37DC9BA9872D45 CRC64;

Query Match 9.8%; Score 82.5; DB 2; Length 389;
Best Local Similarity 24.4%; Pred. No. 4.8;
Matches 47; Conservative 15; Mismatches 74; Indels 57; Gaps 8;

QY 6 ICDSRLVLEKYL-----LEAKAEVNTTGCACGHCLENENITVPTKRVNFYANKRXSKXO 58
DB 75 LASAEILRYLRTPASRIAEDAAVRLVAGAAHMDLKN-----SAS 116
QY 59 OAVEV-----MOGLA-----LTSEAVLNGQALLVNSSQPWEPLQLVKAVSGLSL-- 105
DB 117 QVAEVMGKGTWGVASSDVEITRDAPVLAQAILANK-----HLCVDIKTTSRKTGVH 169
QY 106 -TTLRALGAOKAEKIS-----PPDASAAPLRITITADTFKRLFRVYSNF-----LRGKL 153
DB 170 SKTLERIEGAVVRLSGILEFFPGARPREALRAIGLERFAPPELLINGKIDLDGADLSGJS 229
QY 154 KLTGECARTGDR 166
DB 230 PLYLGITPREADR 242

RESULT 9
ID 08ZDC8 PRELIMINARY; PRT; 323 AA.
AC 08ZDC8:
DT 01-MAR-2002 (TREMBLREL. 20, Created)
DT 01-MAR-2002 (TREMBLREL. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLREL. 21, Last annotation update)
DE Ribonucleoside-diphosphate reductase 2 beta chain (EC 1.17.4.1).
GN NRDP OR YP02648.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxID=632;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / BIOVAR ORIENTALIS;
RX MEDLINE=21470413; PubMed=11586360;
RA Parhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebatia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltham T., Hamlin N., Holtroyd S., Jagsels K., Kariyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).

```

```

DR EMBL: AJ414153; CAC92889.1; -
DR InterPro: IPR000358; RibonucL_redctse.
DR Pfam: PF00268; ribonuc_red_sm; 1.
DR PROSITE: PS00368; RIBORED_SMALL; 1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 323 AA; 36823 MW; 87C21F7BB9B7FD2 CRC64;

Query Match 9.6%; Score 81; DB 16; Length 323;
Best Local Similarity 24.4%; Pred. No. 5.4;
Matches 33; Conservative 19; Mismatches 61; Indels 22; Gaps 5;

QY 38 NITVPTKVNFTAKRNKXKQAVEWQGLALSEAVLNGQALLVNSSQPWEPLQLVH- 96
DB 2 NVKPTIRISALIMWK-IEDDKDLEVMN--RLTSNFWLEKVPKPLSDIPSMATLTPHEQO 58
QY 97 ---KAVSGLSRLTTLRLAQO---KEATSPPDASAAPLRITITADTFKRLFRVYSNFLR 150
DB 59 LTRVFTGLTLDLTONTIGAPALIKDAITPHEAIFSNISFEAVHAYSISFSTL-- 116
QY 151 GKLLKLTGECARTGD 165
DB 117 -----CLMSD 121

RESULT 10
ID 09NB04 PRELIMINARY; PRT; 871 AA.
AC 09NB04:
DT 01-OCT-2000 (TREMBLREL. 15, Created)
DT 01-OCT-2000 (TREMBLREL. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLREL. 20, Last annotation update)
DE Discs LOST.
GN DLT OR BCDNA:LD22238 OR CG12021.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN 11
RP SEQUENCE FROM N.A.
RA Tenentzapf G., Smith C.A., Bhat M., Mogliade J., Tepass U.;
RT "Apical, lateral, and basal polarization cues contribute to the
RT development of the follicular epithelium during Drosophila
RT oogenesis.";
RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF274350; AAF81829.1; -
DR HSSP: O12923; 3pDZ.
DR FlyBase: FBgn0024246; dlt.
DR InterPro: IPR001478; PDZ.
DR Pfam: PF00595; PDZ; 4.
DR SMART: SM00228; PDZ; 4.
DR PROSITE: PSS0106; PDZ; 4.
SQ SEQUENCE 871 AA; 92905 MW; C9A72F6DB4DB5B4D CRC64;

Query Match 9.5%; Score 80.5; DB 5; Length 871;
Best Local Similarity 26.9%; Pred. No. 21;
Matches 50; Conservative 25; Mismatches 72; Indels 39; Gaps 11;

QY 15 YLLEAKE-----AENTTGCACGHCLENENITVPTKRVNFYANK-RNXXKQOAVEV---W 64
DB 335 YVCEKEELSGIFVKVNSPSSA--ADLNGRIYVNDRIIEVDGSLQGSNMQAVELKKGK 392
QY 65 OGIALISEAVLNG-----QALLVN-----SSQPWEPLQLVHDKAVSGLSRLT-----L 108
DB 393 QVAVNLRLEKLVKPKREQLQQAIAANDKLPSAPGTPSPAPMPPTPVATTSATTPPSRSI 452
QY 109 LRALGAOKAEKISPPDASAAP-----LRTITADTF---RKLFRVYSNFLRGKIKLYTGEA 160
DB 453 TRLL--EEBALPAPBEAFMTTPSVTTMTTTLTSSFCAGQLVAVROS-LDGSKRIIPTEV 509
QY 161 CRGDR 166
DB 510 VPLAD 515

```

```

RESULT 11
ID P94873 PRELIMINARY; PRT: 3722 AA.
AC P94873:
DT 01-MAY-1997 (TREMblrel. 03, Created)
DT 01-MAY-1997 (TREMblrel. 03, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Alpha-aminoadipyl-L-cysteine-glycine synthetase.
GN PCBA8.
OS Lysobacter lactamgenus.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Lysobacter.
OC NCBI_TaxID=39596;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VK90.
RX MEDLINE=96275949; PubMed=8737573;
RA Kimura H., Miyashita H., Sumino Y.;
RT "Organization and expression in Pseudomonas putida of the gene cluster
RT involved in cephalosporin biosynthesis from Lysobacter lactamgenus
RT VK90."
RL Appl. Microbiol. Biotechnol. 45:490-501(1996).
DR EMBL: D50308; BAA08846.1; -.
DR HSSP: P14687; IAMU.
DR InterPro: IPR002106; AALRNA_LigaseII.
DR InterPro: IPR000873; AMP-bind.
DR InterPro: IPR001242; Condensatn.
DR InterPro: IPR000977; DNA_Ligase.
DR InterPro: IPR003880; Ppantne_attach.
DR InterPro: IPR000379; Ser_estrs_site.
DR InterPro: IPR001031; Thioesterase.
DR Pfam: PF00501; AMP-binding_3.
DR Pfam: PF00668; Condensation_3.
DR Pfam: PF00550; pp-binding_3.
DR Pfam: PF00975; Thioesterase_1.
DR PROSITE: PS00179; AA_TRNA_LIGASE_II_1; UNKNOWN_1.
DR PROSITE: PS50075; ACP_DOMAIN_3.
DR PROSITE: PS00455; AMP_BINDING_1.
DR PROSITE: PS00697; DNA_LIGASE_A1; UNKNOWN_2.
DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_2.
KW Phosphopantetheine.
SQ SEQUENCE 3722 AA; 411607 MW; 3597B3483463809B CRC64;

Query Match 9.4%; Score 79; DB 2; Length 3722;
Best Local Similarity 22.1%; Pred. No. 1,9e+02;
Matches 34; Conservative 16; Mismatches 52; Indels 52; Gaps 5;

OY 16 LLEAKAEENTTTGCAEHGCSLNENITVPDTKVNFTYAMKRNXSXQAAVEV---WQGLALLS 71
DB 2138 LLEGATTAQTLP-----PLPATYADFVSMOROLSDRLDALFDYWO----- 2178
OY 72 EAVLRGALLVNSQPWEPIQLHYDKAVSGLSRLTTLRALGAQ-----KEAISPDAAS 126
DB 2179 -----RSLAGKQPIQLPLD-----HARPAQDYILGRIEIVDVDAAT 2214
OY 127 AAPLRITTTADTFRKLFRRVYSNFKGLKLYTGEA 160
DB 2215 CDQLRVLAQTTRTSFVSLLAAVYLTLLKAVSQGS 2248

RESULT 12
ID Q920V7 PRELIMINARY; PRT: 383 AA.
AC Q920V7:
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Problem ribonuclease D protein (EC 3.1.26.3).
GN RND OR R01191 OR SMC00622.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

```

```

OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=2136507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Maury D.,
RA Pohl T., Portetle D., Puehler A., Purnelle B., Ransperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL: AL591786; CAC45770.1; -.
DR InterPro: IPR002562; 3_5_exonuclease.
DR InterPro: IPR002121; HRDC.
DR Pfam: PF01612; 3_5_exonuclease; 1.
DR Pfam: PF00570; HRDC; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 383 AA; 42884 MW; 49972055BE17D076 CRC64;

Query Match 9.3%; Score 78.5; DB 16; Length 383;
Best Local Similarity 25.6%; Pred. No. 12;
Matches 45; Conservative 27; Mismatches 57; Indels 47; Gaps 9;

OY 4 RLICDSRYLERYLLEAKEAENITTTGCAEHGCSLNENITVPDTKVNFTYAMKRNXSXQAAVEV 63
DB 233 RLIDDAIYELINQOQPKDAE-----ALGLRTITPK-----GHERSSA----- 269
OY 64 WQGLALLSEAVLRGALLVNSQ-PWEPIQLHYDKAVSGLSRLTTLRALGAQKEAISP 122
DB 270 --GALLI-EAV--SEALAIPEGELPRLPRONHAEAGAAASSELKVLKLISEKEGYAAK 324
OY 123 DAASAPLRITTTAD-----TRKLFRRVYSNFKGLKLYTGE-ACRTQDR 166
DB 325 IIAASDDLRIAGEGADVAALKGWRRELFGMTA-----LKLINGEVALRFDVK 374

RESULT 13
ID Q82ZK3 PRELIMINARY; PRT: 378 AA.
AC Q82ZK3:
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Amidophosphoribosyltransferase (purF).
GN PAF0218.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX PubMed=11792869;
RA Flitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum."
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL: AF009756; AAL62636.1; -.
DR InterPro: IPR000836; PRTtransferase.
DR Pfam: PF00156; Pribosylltran; 1.
KW Transferase; Glycosyltransferase; Complete proteome.
SQ SEQUENCE 378 AA; 40574 MW; 0BE840A6A621A51D CRC64;

Query Match 9.3%; Score 78; DB 17; Length 378;
Best Local Similarity 25.8%; Pred. No. 13;
Matches 42; Conservative 19; Mismatches 56; Indels 46; Gaps 10;

OY 2 PPRICDSRYLERYLLEAKEAENITTTG-CAE-HGCSLNEN-----ITVPDTKVN 47

```

```

DB 173 PREPTEC---AME--LYYASRLDSVIGVEIAEYRAKLGKAKKARADYVICVPTDGLY 227
QY 48 FYAMKRNKXQQAWEVWQGLALLSEAVLRQGLLVNSQWPEPLQAHV-----95
DB 228 YAAW-----AARAGAMWHPAGFVS-TIKKRSALDDEVKERISALQKANYVRHVAAGKRV 281
QY 96 ----DKAVSG--LRSLTTLRALGAQKEAISPPDASAPLRT 132
DB 282 LVVDDSIISGLTRHIAQLIRVAKAKEVHA---ATAAPPLRS 321

RESULT 14
Q8Y76 PRELIMINARY: PRT: 242 AA.
AC Q8Y76:
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE Hypothetical protein A10975.
GN A10975.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-21595285; PubMed-11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriuchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003584; BAB72932.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 242 AA; 27163 MW; 746A46CE10CCF6FC CRC64;

Query Match 9.2%; Score 77.5; DB 16; Length 242;
Best Local Similarity 28.4%; Pred. No. 8.5;
Matches 27; Conservative 14; Mismatches 27; Indels 27; Gaps 4;

QY 66 GLALISEAVL-----RQALLV--NSSQWPEPLQ-----HYDKAVSGLR 103
DB 14 GLCLSSQVLAQTLTLPNLIGFNSNEGKLLTSSREDPPLSMQVTVQVNAVCVGA 73
QY 104 SLTTLRALGAQKEAISPPDASAPLRTTADTF 138
DB 74 STIMVNLISG-----INAPETAQSPYRFTQDNF 103

RESULT 15
Q9SVL5 PRELIMINARY: PRT: 567 AA.
AC Q9SVL5:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Hypothetical 64.1 kDa protein.
GN F18B3.120.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Rieger M., Gabel C., Mueller-Auer S., Schaefer M., Zipp M.,
RA Mewes H.W., Mayer K.F.X., Lemcke K., Schueller C., Quetler F.,
RA Salanoubat M.;
RL Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
RP SEQUENCE FROM N.A.

```

```

RA EU Arabidopsis sequencing project;
RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AL049862; CAB42913.1;
DR InterPro: IPR000210; BRB_P02.
DR InterPro: IPR004249; NPH3.
DR Pfam: PF00651; BRB; 1.
DR Pfam: PF03000; NPH3; 1.
DR SMART; SM00225; BRB; 1.
DR PROSITE; PS50097; BRB; 1.
KW Hypothetical protein.
SQ SEQUENCE 567 AA; 64090 MW; 1A6550B031165D6F CRC64;

```

```

Query Match 9.2%; Score 77.5; DB 10; Length 567;
Best Local Similarity 22.7%; Pred. No. 25;
Matches 47; Conservative 27; Mismatches 76; Indels 57; Gaps 11;

QY 1 APRRLICDSVLERYLLE-----AKEAENITTCGAHCGLNENITVPDRKY-----46
DB 105 SPENLISKI---EKFLSEFVTNVQESIRALKACESVSSLAESLCITEOCIDSIYFQASS 161
QY 47 ---NFYAW-----KRNKXQQAWEW-QGLALLSEAVLRQGLLVNS--QP- 87
DB 162 TDPSSPYGWPINNCGITFTVDRKKQSKDSKTELMFEDLTSPPIFRVILSMKSSVLSPE 221
QY 88 --WEPLQLHVKAVSGL-----RSLTTLRALGAQKEAISPPDASAPLFTTA 135
DB 222 IVERSLITYAKKHIPGISRSSASSSSSSTTIASENQORELLE--TITSDPL--TA 276
QY 136 DTFPRKLFYRYSNPLRGLKLYGCAQR 162
DB 277 TTRSLF-----GLLRATITLNASENCR 299

```

Search completed: January 8, 2003, 05:13:54
 Job time : 56.1267 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 00:55:32 ; Search time 11.5599 Seconds
(without alignments)
595.600 Million cell updates/sec

Title: US-09-813-775c-18

Perfect score: 843
Sequence: 1 APPRLICDSRVLELYLEAK.....NFLRGKMLKLYTGECRTGDR 166

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	825	97.9	193	1	EPO_HUMAN
2	743.5	88.2	192	1	EPO_MACFA
3	741.5	88.0	192	1	EPO_MACMU
4	692	82.1	192	1	EPO_RAT
5	688	81.6	192	1	EPO_FELCA
6	679.5	80.6	192	1	EPO_BOVIN
7	676	80.2	192	1	EPO_MOUSE
8	672.5	79.8	194	1	EPO_SHEEP
9	665	78.9	190	1	EPO_PIG
10	615	73.0	175	1	EPO_CANFA
11	110	13.0	352	1	TPO_CANFA
12	93	11.0	353	1	TPO_HUMAN
13	78	9.3	326	1	TPO_RAT
14	76.5	9.1	543	1	CH60_BARBA
15	76	9.0	356	1	TPO_MOUSE
16	76	9.0	552	1	CH60_COXBU
17	75	8.9	263	1	YH25_DEIRA
18	74	8.8	1089	1	YH25_YEAST
19	73.5	8.7	897	1	YH25_MOUSE
20	73	8.7	353	1	NADA_YERPE
21	71.5	8.5	220	1	YH25_MOUSE
22	71	8.4	547	1	YH25_MOUSE
23	70.5	8.4	381	1	MODD_MYCAV
24	70.5	8.3	3033	1	POLG_HCVU8
25	70	8.3	544	1	CH60_PRAU8
26	69.5	8.2	547	1	CH60_LEGPN
27	69.5	8.2	896	1	EPI5_HUMAN
28	69	8.2	3685	1	DMD_HUMAN
29	68.5	8.1	347	1	NADA_SALIT
30	68.5	8.1	360	1	YH25_MOUSE
31	68.5	8.1	552	1	CH60_PSEST
32	68.5	8.1	907	1	GACS_PSEST
33	68	8.1	542	1	CH65_RHIME

34	68	8.1	548	1	CH60_BUCMP	051832 buchmeta ap
35	67.5	8.0	239	1	GIDB_STRCO	054571 streptomyc
36	67.5	8.0	388	1	TRA6_BURCE	P24575 burkholderi
37	67.5	8.0	778	1	RG12_MOUSE	061193 mus musculu
38	67.5	8.0	1564	1	PDRA_YEAST	P51533 saccharomyc
39	67	7.9	551	1	CH60_YEAST	Q59177 buchmeta ap
40	67	7.9	622	1	FACG_HUMAN	O15287 homo sapien
41	67	7.9	747	1	ATCS_SYNP7	P37279 synechococc
42	67	7.9	1228	1	ATC_PLAFK	008853 plasmodium
43	66.5	7.9	309	1	YBCK_HAEIN	P44298 haemophilus
44	66.5	7.9	511	1	P60_LISCR	001835 listeria gr
45	66.5	7.9	531	1	TRPE_ARTGO	P96556 arthrobacte

ALIGNMENTS

RESULT 1
EPO_HUMAN STANDARD: PRT: 193 AA.
ID EPO_HUMAN
AC P01588; O9UHA0; O9UEZ5; O9UD20;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Erythropoietin precursor (Epoetin).
GN EPO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85137899; PubMed=3838366;
RA Jacobs K., Shoemaker C., Ruderstorf R., Neill S.D., Kaufman R.J.,
RA Mufson A., Seehra J., Jones S.S., Hewick R., Fritsch E.F.,
RA Kawakita M., Shimizu T., Miyake T.,
RT Isolation and characterization of genomic and cDNA clones of human
RT erythropoietin.
RL Nature 313:806-810(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86067948; PubMed=3865178;
RA Lin F.-K., Suggs S., Lin C.-H., Browne J.K., Smalling R., Egrie J.C.,
RA Chen K.K., Fox G.M., Martin F., Stabinsky Z., Badrawi S.M., Lai P.-H.,
RA Goldwasser E.,
RT Cloning and expression of the human erythropoietin gene.
RL Proc. Natl. Acad. Sci. U.S.A. 82:7580-7584(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99018118; PubMed=9799793;
RA Gloeckner G., Scherer S., Schattevoy R., Boright A., Weber J.,
RA Tsui L.-C., Rosenthal A.,
RT Large-scale sequencing of two regions in human chromosome 7q22:
RT analysis of 650 kb of genomic sequence around the EPO and CUTL1 loci
RT reveals 17 genes.
RL Genome Res. 8:1060-1073(1998).
RN [4]
RP SEQUENCE FROM N.A.
RX Ruppert J.L., Hochachka P.W.,
RT Erythropoietin gene sequence in the Quechua, a high altitude native
RT population.
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 58-193 FROM N.A., AND VARIANTS HEPATOCELLULAR CARCINOMA.
RX MEDLINE=93384593; PubMed=8396923;
RA Funakoshi A., Muta H., Baba T., Shimizu S.,
RT Gene expression of mutant erythropoietin in hepatocellular
RT carcinoma.
RL Biochem. Biophys. Res. Commun. 195:717-722(1993).
RN [6]
RP SEQUENCE OF 28-193, AND DISULFIDE BONDS.
RX TISSUE-Urine; MEDLINE=86140080; PubMed=3949763;

RA Lai P.H., Everett R., Wang F.F., Arkawa T., Goldwasser E.;
RP "Structural characterization of human erythropoietin.";
RL J. Biol. Chem. 261:3116-3121(1986).
RN [7]
RP PRELIMINARY SEQUENCE OF 28-57.
RX MEDLINE=84135751; PubMed=6698989;
RA Yanagawa S., Hirade K., Ohnota H., Sasaki R., Chiba H., Ueda M.,
RA Goto M.;
RL "Isolation of human erythropoietin with monoclonal antibodies.";
RN J. Biol. Chem. 259:2707-2710(1984).
RN [8]
RP STRUCTURE OF CARBOHYDRATES.
RX MEDLINE=88153657; PubMed=3346214;
RA Takeuchi M., Takasaki S., Miyazaki H., Kato T., Hoshi S., Koohbe N.,
RA Kobata A.;
RN "Comparative study of the asparagine-linked sugar chains of human
RT erythropoietins purified from urine and the culture medium of
RT recombinant Chinese hamster ovary cells.";
RL J. Biol. Chem. 263:3657-3663(1988).
RN [9]
RP STRUCTURE OF CARBOHYDRATES.
RX MEDLINE=89118279; PubMed=3219367;
RA Sasaki H., Ochi N., Dell A., Fukuda M.;
RN "Site-specific glycosylation of human recombinant erythropoietin:
RT analysis of glycopeptides or peptides at each glycosylation site by
RT fast atom bombardment mass spectrometry.";
RL Biochemistry 27:8618-8626(1988).
RN [10]
RP STRUCTURE OF CARBOHYDRATES.
RX MEDLINE=92314463; PubMed=1820196;
RA Takeuchi M., Kobata A.;
RN "Structures and functional roles of the sugar chains of human
RT erythropoietins.";
RL Glycobiology 1:337-346(1991).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=98445092; PubMed=9774108;
RA Syed R.S., Reid S.W., Li C., Cheeham J.C., Aoki K.H., Liu B.,
RA Zhan H., Osslund T.D., Chirino A.J., Zhang J., Flier-Moore T.,
RA Elliott S., Stoney R., Katz B.A., Matthews D.J., Wendoloski J.J.,
RA Eglye J., Stroud R.M.;
RN "Efficiency of signalling through cytokine receptors depends
RT critically on receptor orientation.";
RL Nature 395:511-516(1998).
CC -I- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC -I- PHARMACEUTICAL: Used for the treatment of anemia. Available under
CC the names Epogen (Amgen), Epogin (Chugai), Epomax (Eliane), Eprex
CC (Janssen-Cilag), Neorecormon or Recormon (Boehr), and Procrit
CC (Ortho Biotech). Variations in the glycosylation pattern of Epo
CC distinguishes these products. Epogen, Epogin, Eprex and Procrit
CC are genetically known as epoetin alfa, Neorecormon and Recormon as
CC epoetin beta and Epomax as epoetin omega.
CC -I- SIMILARITY: BELONGS TO THE Epo / TPO FAMILY.
CC -I- DATABASE: NAME=R&D Systems' cytokine source book: Epo:
CC WWW="http://www.rndsystems.com/asp/9_sitbuilder.asp?bodyId=197".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X02158; CAA26095.1; -
DR EMBL: X02157; CAA26094.1; -
DR EMBL: M13119; AAA52400.1; -
DR EMBL: AF053356; AAC8791.1; -

Query	Match	Best Local Similarity	Score	DB 1:	Length	DB 2:
Matches	163:	Conservative	0:	Mismatches	4:	Indels
QY 1	APPRIICDSRYLERYLLLEAKAEENITGGCAHCSLNENITVPDTRKVNRYANKRNXSQA	60				
Db 28	APPRIICDSRYLERYLLLEAKAEENITGGCAHCSLNENITVPDTRKVNRYANKRNXSQA	87				
QY 61	VEVNOGIALISEAVLYRGQALLVNSSQPEWPEQLQTHYDKAVSGISRLTTLRLAQAQKAIS	120				
Db 88	VEVNOGIALISEAVLYRGQALLVNSSQPEWPEQLQTHYDKAVSGISRLTTLRLAQAQKAIS	147				
QY 121	PPDASAPLFTTTADTFPRKLFRRVYSNFRGKRLKLYGEGACRTGDR	166				
Db 148	PPDASAPLFTTTADTFPRKLFRRVYSNFRGKRLKLYGEGACRTGDR	193				
RESULT 2						
EPD_MACPA	STANDARD;	PRT;	192 AA.			
AC P07865;						
DT 01-AUG-1988 (Rel. 08, Created)						
DT 01-AUG-1988 (Rel. 08, Last sequence update)						
DT 16-OCT-2001 (Rel. 40, Last annotation update)						
DE Erythropoietin precursor						

GN EPO.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 NCBI_TaxID=9541;
 RN NCBI_TaxID=9541;
 RP SEQUENCE FROM N.A.
 RA MEDLINE=87055236; PubMed=2877922;
 RX Lin F.-K., Lin C.-H., Lai P.-H., Browne J.K., Egrie J.C., Smalling R.,
 RA Fox G.M., Chen K.K., Castro M., Suggs S.;
 RT "Monkey erythropoietin gene: cloning, expression and comparison with
 the human erythropoietin gene.";
 RL Gene 44:201-209(1986).
 CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
 CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
 CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
 CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
 CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL_M18189; AAA36841.1; -;
 DR PIR; J00173; J00173;
 DR HSSP; P01588; ICN4.
 DR InterPro: IPR001323; EPO_TPO.
 DR Pfam: PF00758; EPO_TPO.1.
 DR PRINTS: PR00272; ERYTHROPTN.
 DR PROSITE: PS00817; EPO_TPO.1.
 KW Erythrocyte maturation; Glycoprotein; Hormone; Signal.
 FT SIGNAL 1 27
 FT CHAIN 1 192
 FT DISULFID 34 187
 FT DISULFID 56 60
 FT CARBOHYD 51 51
 FT CARBOHYD 65 65
 FT CARBOHYD 110 110
 FT CARBOHYD 152 152
 FT CARBOHYD 192 AA; 21113 MW; E8A900F42AD4522 CRC64;
 SQ SEQUENCE
 Query Match 88.2%; Score 743.5; DB 1; Length 192;
 Best Local Similarity 89.2%; Pred. No. 2.2e-68;
 Matches 148; Conservative 7; Mismatches 10; Indels 1; Gaps 1;
 Oy 1 APRRLCDSDVRLRYLLLEAKENITTCGAHCSLNNITVPDKRVFYAMKRNKXSOQA 60
 Db 28 APRRLCDSDVRLRYLLLEAKENITTCGAHCSLNNITVPDKRVFYAMKRNKXSOQA 87
 Oy 61 VEVWOGTALLSEAVLRGQAVLVNSSQPEPQLQHVDAVSGLSRLTTLRLALGQKRAIS 120
 Db 88 VEVWOGTALLSEAVLRGQAVLVNSSQPEPQLQHVDAVSGLSRLTTLRLALGQKRAIS 146
 Oy 121 PPDAASAAPLRTITADTFKRLFRVYSNPLRGKLTLYGCAACRTGDR 166
 Db 147 LPDAASAAPLRTITADTFKRLFRVYSNPLRGKLTLYGCAACRTGDR 192
 RESULT 3
 EPO_MACHU
 ID EPO_MACHU STANDARD; PRT; 192 AA.
 AC Q28513;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Erythropoietin precursor.
 GN EPO.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 NCBI_TaxID=9544;
 RN NCBI_TaxID=9544;
 RP SEQUENCE FROM N.A.
 RA MEDLINE=93372347; PubMed=8364201;
 RX Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
 RA Czelusniak J., Goodman M., Bunn H.F.;
 RT "Erythropoietin structure-function relationships: high degree of
 RT sequence homology among mammals.";
 RL Blood 82:1507-1516(1993).
 CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
 CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
 CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
 CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
 CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L10609; AAA36842.1; -;
 DR HSSP; P01588; ICN4.
 DR InterPro: IPR001323; EPO_TPO.
 DR Pfam: PF00758; EPO_TPO.1.
 DR PRINTS: PR00272; ERYTHROPTN.
 DR PROSITE: PS00817; EPO_TPO.1.
 KW Erythrocyte maturation; Glycoprotein; Hormone; Signal.
 FT SIGNAL 1 27
 FT CHAIN 1 192
 FT DISULFID 34 187
 FT DISULFID 56 60
 FT CARBOHYD 51 51
 FT CARBOHYD 65 65
 FT CARBOHYD 110 110
 FT CARBOHYD 152 152
 FT CARBOHYD 192 AA; 21081 MW; 275560A264628CD1 CRC64;
 SQ SEQUENCE
 Query Match 88.0%; Score 741.5; DB 1; Length 192;
 Best Local Similarity 88.6%; Pred. No. 3.5e-68;
 Matches 147; Conservative 8; Mismatches 10; Indels 1; Gaps 1;
 Oy 1 APRRLCDSDVRLRYLLLEAKENITTCGAHCSLNNITVPDKRVFYAMKRNKXSOQA 60
 Db 28 APRRLCDSDVRLRYLLLEAKENITTCGAHCSLNNITVPDKRVFYAMKRNKXSOQA 87
 Oy 61 VEVWOGTALLSEAVLRGQAVLVNSSQPEPQLQHVDAVSGLSRLTTLRLALGQKRAIS 120
 Db 88 VEVWOGTALLSEAVLRGQAVLVNSSQPEPQLQHVDAVSGLSRLTTLRLALGQKRAIS 146
 Oy 121 PPDAASAAPLRTITADTFKRLFRVYSNPLRGKLTLYGCAACRTGDR 166
 Db 147 LPDAASAAPLRTITADTFKRLFRVYSNPLRGKLTLYGCAACRTGDR 192
 RESULT 4
 EPO_RAT
 ID EPO_RAT STANDARD; PRT; 192 AA.
 AC P29676; P70504;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)

16-OCT-2001 (Rel. 40, last annotation update)
 Erythropoietin precursor.
 GN EPO.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI_TaxId=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mistar; TISSUE=Kidney;
 RX MEDLINE=93042015; PubMed=1420369;
 RA Nagao M., Suga H., Okano M., Masuda S., Narita H., Ikura K.,
 RA Sasaki R.;
 RT "Nucleotide sequence of rat erythropoietin";
 RL Blochim. Biophys. Acta 1171:99-102(1992).
 RN (2)
 RP SEQUENCE OF 4-192 FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Kidney;
 RX MEDLINE=93372347; PubMed=8364201;
 RA Wen D., Boissel J.P.R., Tracy T.E., Mulcahy L.S., Czelusniak J.,
 RA Goodman M., Bunn H.F.;
 RT "Erythropoietin structure-function relationships: high degree of
 RT sequence homology among mammals";
 RL Blood 82:1507-1516(1993).
 CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
 CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
 CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
 CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
 CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
 CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: D10763; BAA01593.1; -;
 DR EMBL: L10608; AAA41126.1; -;
 DR PIR: S28148; S28148.
 DR HSSP: P01588; ICN4.
 DR InterPro: IPR001323; EPO_TPO.
 DR InterPro: IPR003013; Erythropo.
 DR Pfam: PF00758; EPO_TPO.1.
 DR PRINTS: PR00272; ERYTHROPTN.
 DR PROSITE: PS00817; EPO_TPO.1.
 KM Erythrocyte maturation; Glycoprotein; Hormone; Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 192
 FT DISULFID 33 187
 FT CARBOHYD 50 50
 FT CARBOHYD 64 64
 FT CARBOHYD 109 109
 SEQUENCE 192 AA: 21286 MW: 35632737 E7D2443 CRC64;
 Query Match 81.1%; Score 692; DB 1; Length 192;
 Best Local Similarity 81.3%; Pred. No. 3,8e-63;
 Matches 135; Conservative 12; Mismatches 19; Indels 0; Gaps 0;
 QY 1 APRRLICDSRVLYRLLEAKKAEINITTCGAECSCSINENITVDTKYNFYAMRNKNSXQA 60
 DB 27 APRRLICDSRVLYRLLEAKKAEINITTCGAECSCSINENITVDTKYNFYAMRNKNSXQA 86
 QY 61 VEWVGGLALLSAVLRGQALLVNSSQWPEPDLQHYDKAVSGIRSLTTLRLAAGAKKELAS 120
 DB 87 VEWVGGLALLSAVLRGQALLVNSSQWPEPDLQHYDKAVSGIRSLTTLRLAAGAKKELAS 146
 QY 121 PPDAAASAPLRTITADTFKRLFRVYSNPLRGKLYGKLYTGACRTGR 166
 DB 147 PPDAAQAPLRTITADTFKRLFRVYSNPLRGKLYGKLYTGACRTGR 192

RESULT 5
 EPO_FELCA
 ID EPO_FELCA STANDARD: PRT: 192 AA.
 AC P33708;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-OCT-1996 (Rel. 34, last sequence update)
 DT 16-OCT-2001 (Rel. 40, last annotation update)
 DE Erythropoietin precursor.
 GN EPO.
 OS Felis silvestris catus (Cat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
 CC NCBI_TaxId=9685;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Goodman R.E., Bell R.G.;
 RT Submitted (NOV-1993) to the EMBL/Genbank/DBJ databases.
 RN (2)
 RP SEQUENCE OF 5-192 FROM N.A.
 RX MEDLINE=93372347; PubMed=8364201;
 RA Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
 RA Czelusniak J., Goodman M., Bunn H.F.;
 RT "Erythropoietin structure-function relationships: high degree of
 RT sequence homology among mammals";
 RL Blood 82:1507-1516(1993).
 CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
 CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
 CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
 CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
 CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
 CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U00685; AAA18282.1; -;
 DR EMBL: L10606; AAA30807.1; -;
 DR HSSP: P01588; ICN4.
 DR InterPro: IPR001323; EPO_TPO.
 DR InterPro: IPR003013; Erythropo.
 DR Pfam: PF00758; EPO_TPO.1.
 DR PRINTS: PR00272; ERYTHROPTN.
 DR PROSITE: PS00817; EPO_TPO.1.
 KM Erythrocyte maturation; Glycoprotein; Hormone; Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 192
 FT DISULFID 33 187
 FT CARBOHYD 50 50
 FT CARBOHYD 64 64
 FT CARBOHYD 109 109
 FT CONFLICT 44 44
 SEQUENCE 192 AA: 20914 MW: 61C5EA0F5E937293 CRC64;
 Query Match 81.6%; Score 688; DB 1; Length 192;
 Best Local Similarity 81.9%; Pred. No. 9,6e-63;
 Matches 136; Conservative 8; Mismatches 22; Indels 0; Gaps 0;
 QY 1 APRRLICDSRVLYRLLEAKKAEINITTCGAECSCSINENITVDTKYNFYAMRNKNSXQA 60
 DB 27 APRRLICDSRVLYRLLEAKKAEINITTCGAECSCSINENITVDTKYNFYAMRNKNSXQA 86
 QY 61 VEWVGGLALLSAVLRGQALLVNSSQWPEPDLQHYDKAVSGIRSLTTLRLAAGAKKELAS 120

[illegible]

```

Db      86 LEVWGQALLLSFAIRIRGALLANASQPEALRIHYDKAVSGILRSITSLRLACQAKEXAIS 145
QY      121 PPDAA-SAAPLRTITADTFRKLEFRVYSNFKLGKTKLTYTGACRTGD 166
Db      146 LPDAPSAAPLRAFTVDALSKLFRTYSNFKLGKTLTYTGACRRDDR 192

RESULT 7
EPO_MOUSE
ID      EPO_MOUSE      STANDARD;      PRT;      192 AA.
AC      P07321;
DT      01-APR-1988 (Rel. 07, Created)
DT      01-APR-1988 (Rel. 07, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Erythropoietin precursor.
GN      EPO.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=87039105; PubMed=3773894;
RA      Shoemaker C.B., Mitsock L.D.;
RT      "Murine erythropoietin gene: cloning, expression, and human gene
RT      homology.";
RL      Mol. Cell. Biol. 6:849-858(1986).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=87039104; PubMed=3022133;
RA      McDonald J.D., Lin F.-K., Goldwasser E.;
RT      "Cloning, sequencing, and evolutionary analysis of the mouse
RT      erythropoietin gene.";
RL      Mol. Cell. Biol. 6:842-848(1986).
RN      [3]
RP      SEQUENCE FROM N.A.
RX      STRAIN=129/SV;
RA      MEDLINE=21138439; PubMed=11239002;
RA      Wilson M.D., Riemer C., Martindale D.W., Schnupf P., Borlight A.P.,
RA      Cheung T.L., Hardy D.M., Schwartz S., Scherer S.W., Tsui L.-C.,
RA      Miller W., Koop B.F.;
RT      "Comparative analysis of the gene-dense ACHE/TFR2 region on human
RT      chromosome 7q22 with the orthologous region on mouse chromosome 5.";
RL      Nucleic Acids Res. 29:1352-1365(2001).
RN      [4]
RP      SEQUENCE OF 1-52 FROM N.A.
RX      STRAIN=ICRW;
RX      MEDLINE=98030528; PubMed=9365246;
RA      Chretien S., Duprez V., Maouche L., Gisselbrecht S., Mayeux P.,
RA      Lacombe C.;
RT      "Abnormal erythropoietin (Epo) gene expression in the murine
RT      erythroleukemia Iw32 cells results from a rearrangement between the
RT      G-protein beta2 subunit gene and the Epo gene.";
RL      Oncogene 15:1195-1199(1997).
RN      [5]
RP      PUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
RP      REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
RP      PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
RN      [6]
RP      SUBCELLULAR LOCATION: Secreted.
RN      [7]
RP      TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
RN      [8]
RP      AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
RN      [9]
RP      -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
RN      [10]
RP      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
RP      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
RP      the European Bioinformatics Institute. There are no restrictions on its
RP      use by non-profit institutions as long as its content is in no way
RP      modified and this statement is not removed. Usage by and for commercial
RP      entities requires a license agreement (See http://www.isb-sib.ch/announce/
RP      or send an email to license@isb-sib.ch).
RN      [11]
EMBL; M12482; AAA37568.1; -
EMBL; M12930; AAA37570.1; -

```

EMBL: AF312033; AK28825.1; -
 DR EMBL: Y11971; CA472707.1; -
 DR PIR: A24901; A24901.
 DR PIR: A24902; A24902.
 DR HSP: P01588; ICN4.
 DR MGD: MGI:95407; EPO.
 DR InterPro: IPR001323; EPO_TPO.
 DR InterPro: IPR003013; Erythropo.
 DR Pfam: PF00758; EPO_TPO.1.
 DR PRINTS: PR00272; ERYTHROPTN.
 DR PROSITE: PS00817; EPO_TPO.1.
 KW Erythrocyte maturation; Glycoprotein; Hormone; signal.
 FT SIGNAL 1 26
 FT CHAIN 27 192
 FT DISULFID 33 187 BY SIMILARITY.
 FT CARBOHYD 50 50 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
 FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
 FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
 SQ SEQUENCE 192 AA: 21365 MW: 6594E214E0DEF2E CRC64;

Query Match 80.2%; Score 676; DB 1; Length 192;
 Best Local Similarity 78.3%; Pred. No. 1.6e-61;
 Matches 130; Conservative 14; Mismatches 22; Indels 0; Gaps 0;

OY 1 APRRLICDSRVLYERLYLEAKENITTCGAHCISNENITVPDTKYNFYAMKRNKXQQA 60
 DB 27 APRRLICDSRVLYERLYLEAKENITTCGAHCISNENITVPDTKYNFYAMKRNKXQQA 86
 OY 61 VEWMOGLALISPAVLRGQALLVNSSQWPEPLQHVDAKAVSGIRSLTTLRALGAQKEAS 120
 DB 87 LEWMOGLALISPAVLRGQALLVNSSQWPEPLQHVDAKAVSGIRSLTTLRALGAQKEAS 146
 OY 121 PPDAASAAPLRTITADTFRKLFYVSNFLRGKLLKLYTGACRTGDR 166
 DB 147 PDDTTPAPALRTITADTFRKLFYVSNFLRGKLLKLYTGACRTGDR 192

RESULT 8

EPO_SHEEP STANDARD; PRT; 194 AA.

AC P33709; Q28572;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Erythropoietin precursor.
 GN EPO.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Kidney;
 RX MEDLINE=93351736; PubMed=8349021;
 RA Fu P., Evans B., Lim G.B., Moritz K., Wintour M.E.;
 RT "The sheep erythropoietin gene: molecular cloning and effect of
 RT hemorrhage on plasma erythropoietin and renal/liver messenger RNA in
 RT adult sheep.";
 RL Mol. Cell. Endocrinol. 93:107-116(1993).
 RL [2]
 RP SEQUENCE OF 4-194 FROM N.A.
 RC TISSUE-Kidney;
 RX MEDLINE=93372347; PubMed=8364201;
 RA Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
 RT Czelusniak J., Goodman M., Bunn H.F.;
 RT "Erythropoietin structure-function relationships: high degree of
 RT sequence homology among mammals.";
 RL Blood 82:1507-1516(1993).
 CC - FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
 CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
 CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
 CC - SUBCELLULAR LOCATION: Secreted.

CC - TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
 CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
 CC - SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>
 CC or send an email to license@isb.ch).
 CC -----

DR EMBL: Z24681; CA80848.1; -
 DR EMBL: L10610; AAA31518.1; -
 DR HSP: P01588; ICN4.
 DR InterPro: IPR001323; EPO_TPO.
 DR InterPro: IPR003013; Erythropo.
 DR Pfam: PF00758; EPO_TPO.1.
 DR PRINTS: PR00272; ERYTHROPTN.
 DR PROSITE: PS00817; EPO_TPO.1.
 KW Erythrocyte maturation; Glycoprotein; Hormone; signal.
 FT SIGNAL 1 27
 FT CHAIN 28 194
 FT DISULFID 34 189 BY SIMILARITY.
 FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 16 16 F -> L (IN REF. 2).
 FT CONFLICT 108 108 L -> P (IN REF. 2).
 SQ SEQUENCE 194 AA: 21335 MW: C025AAB0528131A9 CRC64;

Query Match 79.8%; Score 672.5; DB 1; Length 194;
 Best Local Similarity 80.2%; Pred. No. 3.7e-61;
 Matches 134; Conservative 9; Mismatches 23; Indels 1; Gaps 1;

OY 1 APRRLICDSRVLYERLYLEAKENITTCGAHCISNENITVPDTKYNFYAMKRNKXQQA 60
 DB 28 APRRLICDSRVLYERLYLEAKENITTCGAHCISNENITVPDTKYNFYAMKRNKXQQA 87
 OY 61 VEWMOGLALISPAVLRGQALLVNSSQWPEPLQHVDAKAVSGIRSLTTLRALGAQKEAS 120
 DB 88 LEWMOGLALISPAVLRGQALLVNSSQWPEPLQHVDAKAVSGIRSLTTLRALGAQKEAIP 147
 OY 121 PPDAASAAPLRTITADTFRKLFYVSNFLRGKLLKLYTGACRTGDR 166
 DB 148 LPDAPPSAAPLRTITADTFRKLFYVSNFLRGKLLKLYTGACRTGDR 194

RESULT 9

EPO_PIG STANDARD; PRT; 190 AA.

AC P49157;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Erythropoietin precursor (Fragment).
 GN EPO.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Kidney;
 RX MEDLINE=93372347; PubMed=8364201;
 RA Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
 RT Czelusniak J., Goodman M., Bunn H.F.;
 RT "Erythropoietin structure-function relationships: high degree of
 RT sequence homology among mammals.";
 RL Blood 82:1507-1516(1993).
 CC - FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
 CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A

CC	-1	TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC	-1	AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC	-1	SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC	-1	-----
CC	-1	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	-1	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	-1	the European Bioinformatics Institute. There are no restrictions on its
CC	-1	use by non-profit institutions as long as its content is in no way
CC	-1	modified and this statement is not removed. Usage by and for commercial
CC	-1	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	-1	or send an email to license@isb-sib.ch).
CC	-1	-----
DR	EMBL	L13027; AAA30842.1; -
DR	HSSP	P01588; ICN4.
DR	InterPro	IPR001323; EPO_TPO.
DR	Pfam	PF00758; EPO_TPO.1.
DR	PROSITE	PS00817; EPO_TPO.1.
KW	Erythrocyte maturation	; Glycoprotein; Hormone; signal.
FT	NON_TER	1
FT	SIGNAL	<1 22
FT	CHAIN	23 >175
FT	DISULFID	29 >175
FT	DISULFID	51 55
FT	CARBOHYD	46 46
FT	CARBOHYD	60 60
FT	CARBOHYD	105 105
FT	NON_TER	175 175
SQ	SEQUENCE	175 AA; 19193 MW; B504F8DE6676BF4 CRC64;

Query Match	73.0%	Score 615;	DB 1;	Length 175;
Best Local Similarity	79.1%	Pred. No. 2.2e-55;		
Matches 121;	Conservative 12;	Mismatches 20;	Indels 0;	Gaps 0;
Oy	1	APPRICDSRVLEKRLTAKKAEKENTTCQAEHCSENENITVPDTRVYNKAMKRNKXSQQA 60		
Db	23	APPRICDSRVLEKRLTAKKAEKENTTCQAEHCSENENITVPDTRVYNKAMKRNKXSQQA 82		
Oy	61	VEWVGALLSEAVYLRGQALLVYNSQPEPQLDHYDKAVSGHSLTTLRLAAGAKKAS 120		
Db	83	LEWVGALLSEAVYLRGQALLVYNSQPEPQLDHYDKAVSGHSLTTLRLAAGAKKAS 142		
Oy	121	PPDASAPLPRTITADTRFKLFRVYSNFLRGKL 153		
Db	143	LPENASAPLPRTITADTRFKLFRVYSNFLRGKL 175		

RESULT 11				
TPO_CANFA				
ID	TPO_CANFA	STANDARD:	PRT:	352 AA.
AC	P42705;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Thrombopoietin precursor (Megakaryocyte colony stimulating factor)			
DE	(C-MPL ligand) (ML) (Megakaryocyte growth and development factor)			
DE	(MCDP).			
GN	THPO OR TPO.			
OS	Canis familiaris (Dog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.			
OX	NCBI_TaxID=9615;			
RP	[1]			
RN	SEQUENCE FROM N.A., AND SEQUENCE OF 24-44.			
RC	TISSUE=Kidney;			
RX	MEDLINE=94291201; PubMed=8020099;			
RA	Battley T.D., Bogenberger J., Hunt P., Li Y.-S., Lu H.S., Martin F.,			
RA	Chang M.-S., Samal B.B., Nichol J.L., Swift S., Johnson M.J.,			
RA	Hsu R.-Y., Parker V.P., Suggs S., Skrine J.D., Merewether L.A.,			
RA	Clogson C., Hsu E., Horkm M.M., Hornkohl A., Choi E., Pangelman M.,			
RA	Sun Y., Mar V., McNich J., Simonet L., Jacobsen F., Xie C.,			
RA	Shutter J., Chute H., Basu R., Selander L., Trollinger D., Siew L.,			
RA	Padilla D., Traill G., Elliott G., Izumi R., Covey T., Crouse J.,			
RA	Garcia A., Xu W., del Castillo J., Biron J., Cole S., Hu M.C.-T.,			

RA Pacific R., Ponting I., Saris C., Wen D., Yung Y.P., Lin H.,
 RA Bosseman R.A.;
 RT "Identification and cloning of a megakaryocyte growth and development
 RT factor that is a ligand for the cytokine receptor Mpl.";
 RL Cell 77:1117-1117(1994).
 CC -1- FUNCTION: LINEAGE-SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION
 CC AND MATURATION OF MEGAKARYOCYTES FROM THEIR COMMITTED PROGENITOR
 CC CELLS. IT ACTS AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT
 CC MAY BE THE MAJOR PHYSIOLOGICAL REGULATOR OF CIRCULATING PLATELETS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DOMAIN: TWO-DOMAIN STRUCTURE WITH AN ERYTHROPOIETIN-LIKE N-
 CC TERMINAL AND A SER/PRO/THR-RICH C-TERMINAL.
 CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
 DR InterPro: IPR001323; EPO_TPO.
 DR InterPro: IPR003978; thrombopoietin.
 DR Pfam: PF00758; EPO_TPO; 1.
 DR PRINTS: PF01485; THROMBOPTN.
 DR PROSITE: PS00817; EPO_TPO; 1.
 KM Cytokine; Glycoprotein; Hormone; Signal.
 FT SIGNAL 1 23
 FT CHAIN 24 352 THROMBOPOIETIN.
 FT DISULFID 28 172 POTENTIAL.
 FT DISULFID 50 106 POTENTIAL.
 FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 352 AA; 73641 MW; 024F3B41B061FBDB8 CRC64;
 Query Match 13.0%; Score 110; DB 1; Length 352;
 Best Local Similarity 24.0%; Pred. NO. 0.0011;
 Matches 40; Conservative 22; Mismatches 63; Indels 42; Gaps 5;
 QY 1 APPRLICDSRVLEKLEKKEKNTTGCAGKSLNENTVVDVNFAMKRNKXQQA 60
 DB 24 APP--ACDPRLNKMRLSHVSRISQCPDTPPLTPVLPADVSLGEMTQREQTKA 81
 QY 61 VFWOGLALISEAVL--RQALLVNNSQWPEPLQLHVDKAVSLSTLTLLALGNOKEA 118
 DB 82 ODVCAVALLDGVLAARQL-----GPSCLSLSLQLQSLGQVNL 120
 QY 119 I-----SPDAAASAPRLRTTADTFERKLPFRVYSNFKRLK 154
 DB 121 LLAGLGLGLGTQLPPG-----RTTHKDPNALFLSFGQLLRKGYR 161
 RESULT 12
 ID TPO_HUMAN STANDARD; PRT; 353 AA.
 AC P40225; Q13020; Q15790; Q15791; Q15792;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Thrombopoietin precursor (Megakaryocyte colony stimulating factor)
 DE (Myeloproliferative leukemia virus oncogene ligand) (C-mpl ligand)
 DE (Mpl) (Megakaryocyte growth and development factor) (MGDF).
 GN THPO.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Fetal liver;
 RX MEDLINE=94261202; PubMed=8202154;
 RA de Sauvage F.J., Haas P.E., Spencer S.D., Malloy B.E., Gurney A.L.,
 RA Spencer S.A., Darbonne W.C., Henzel W.J., Wong S.C., Kuang W.-J.,
 RA Oles K.J., Hultgren B., Solberg L.A. Jr., Goeddel D.V., Eaton D.L.;
 RT "Stimulation of megakaryocytopoiesis and thrombopoiesis by the c-Mpl
 RT ligand.";

RL Nature 369:533-538(1994).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Fetal liver;
 RX MEDLINE=94291202; PubMed=8020099;
 RA Bartley T.D., Bogenberger J., Hunt P., Li Y.-S., Lu H.S., Martin F.,
 RA Chang M.-S., Samal B.B., Nichol J.L., Swift S., Johnson M.J.,
 RA Hsu R.-Y., Parker V.P., Sugis S., Skrine J.D., Merewether L.A.,
 RA Clogson C., Hsu E., Hoxom M.M., Hornkohl A., Choi E., Pangellian M.,
 RA Sun Y., Mar V., McNich J., Simonet L., Jacobsen F., Xie C.,
 RA Shutter J., Chute H., Basu R., Selander L., Trollinger D., Siu L.,
 RA Padilla D., Trail G., Elliott G., Izumi R., Covey T., Crouse J.,
 RA Garcia A., Xu W., del Castillo J., Biron J., Cole S., Hu M.C.-T.,
 RA Pacific R., Ponting I., Saris C., Wen D., Yung Y.P., Lin H.,
 RA Bosseman R.A.;
 RT "Identification and cloning of a megakaryocyte growth and development
 RT factor that is a ligand for the cytokine receptor Mpl.";
 RL Cell 77:1117-1124(1994).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=95108091; PubMed=7809166;
 RA Foster D.C., Sprecher C.A., Grant F.J., Kramer J.M., Kuipper J.L.,
 RA Holly R.D., Whitmore T.E., Heipel M.D., Bell L.A.N., Ching A.F.,
 RA McGlane V., Hart C., O'Hara P.J., Lok S.,
 RT "Human thrombopoietin: gene structure, cDNA sequence, expression, and
 RT chromosomal localization.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:13023-13027(1994).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=95010765; PubMed=7926023;
 RA Sohma Y., Akahori H., Seki N., Hori T.-A., Ogami K., Kawamura K.,
 RA Miyazaki H.;
 RT "Molecular cloning and chromosomal localization of the human
 RT thrombopoietin gene.";
 RL FEBS Lett. 353:57-61(1994).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RX MEDLINE=95152076; PubMed=7849319;
 RA Gurney A.L., Kuang W.-J., Xie M.-H., Malloy B.E., Eaton D.L.,
 RA de Sauvage F.J.;
 RT "Genomic structure, chromosomal localization, and conserved
 RT alternative splice forms of thrombopoietin.";
 RL Blood 85:981-988(1995).
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Liver;
 RX MEDLINE=96015174; PubMed=8537317;
 RA Kato T., Ogami K., Shimada Y., Iwamatsu A., Sohma Y., Akahori H.,
 RA Hori K., Kokubo A., Kudo Y., Maeda E., Kobayashi K., Ohashi H.,
 RA Ozawa T., Inoue H., Kawamura K., Miyazaki H.;
 RT "Purification and characterization of thrombopoietin.";
 RL J. Biochem. 118:229-236(1995).
 RN [7]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Placenta;
 RX MEDLINE=95122483; PubMed=7822271;
 RA Chang M., McNich J., Basu R., Shutter J., Hsu R., Perkins C., Mar V.,
 RA Sugs S., Welcher A., Li L., Lu H., Bartley T., Hunt P., Martin F.,
 RA Samal B., Bogenberger J.;
 RT "Cloning and characterization of the human megakaryocyte growth and
 RT development factor (MGDF) gene.";
 RL J. Biol. Chem. 270:511-514(1995).
 RN [8]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
 RA Im S.H., Lee W.S., Chung K.H.;
 RT "Cloning and sequencing of human thrombopoietin.";
 RL Submitted (May-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: LINEAGE-SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION
 CC AND MATURATION OF MEGAKARYOCYTES FROM THEIR COMMITTED PROGENITOR
 CC CELLS. IT ACTS AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT
 CC MAY BE THE MAJOR PHYSIOLOGICAL REGULATOR OF CIRCULATING PLATELETS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2/TPO-2 AND


```

CC 3/TUNCATED: ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- DOMAIN: TWO-DOMAIN STRUCTURE WITH AN ERYTHROPOIETIN-LIKE N-
CC TERMINAL AND A SER/PRO/THR-RICH C-TERMINAL.
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC -1- DATABASE: NAME-REID systems' cytokine source book: TPO:
CC WWW="http://www.rndsystems.com/asp/g_sitebuilder.asp?bodyid=225".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L33410; AAA59857.1; -
DR EMBL: U11025; AAA50553.1; -
DR EMBL: L36051; AAC37568.1; -
DR EMBL: L36052; AAC37566.1; -
DR EMBL: D32046; BAA06807.1; -
DR EMBL: S76771; AAB33390.1; -
DR EMBL: D32047; BAA21930.1; -
DR EMBL: U59493; AAB03392.1; -
DR EMBL: U59494; AAB03393.1; -
DR EMBL: U59495; AAB03394.1; -
DR EMBL: U17071; AAA74083.1; -
DR PIR: S45331; S45331.
DR Gene: HGNC:11795; THPO.
DR MIM: 600044; -
DR InterPro: IPR001333; EPO_TPO.
DR InterPro: IPR003978; Thrombopoietin.
DR Pfam: PF00758; EPO_TPO; 1.
DR PRINTS: PR01485; THROMBOPTN.
DR PROSITE: PS00817; EPO_TPO; 1.
KM Cytokine; Glycoprotein; Hormone; Signal; Alternative splicing;
KW Polymorphism.
FT SIGNAL 1 21
FT CHAIN 22 353
FT DISULFID 28 172
FT CARBOHYD 50 106
FT CARBOHYD 197 197
FT CARBOHYD 206 206
FT CARBOHYD 234 234
FT CARBOHYD 255 255
FT CARBOHYD 340 340
FT CARBOHYD 348 348
FT VARSPLIC 133 136
FT VARSPLIC 160 198
FT VARIANT 14 14
FT VARIANT 14 14
FT VARIANT 116 116
FT CONFLICT 46 46
FT CONFLICT 76 76
FT CONFLICT 113 113
FT CONFLICT 131 131
FT CONFLICT 277 277
FT CONFLICT 346 346
SQ SEQUENCE 353 AA: 37822 MW: F0AB5449B72E5526 CRC64;

Query Match 11.0%; Score 93; DB 1; Length 353;
Best Local Similarity 26.3%; Pred. No. 0.057;
Matches 41; Conservative 20; Mismatches 75; Indels 20; Gaps 5;

```

```

DB 133 -LPPQG-----RTYAKDPAIFLSPFHLLRQKVR 161
RESULT 13
TPO_RAT
ID TPO_RAT STANDARD; PRT; 326 AA.
AC P49745;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Thrombopoietin precursor.
GN THPO.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID:10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=9531639; Pubmed=7607561;
RA Ogami K., Shimada Y., Sohma Y., Akahori H., Kato T., Kawamura K.,
RA Miyazaki H.:
RT "The sequence of a rat cDNA encoding thrombopoietin."
RL Gene 158:309-310(1995).
CC -1- FUNCTION: SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION
CC AND MATURATION OF MEGAKARYOCYTES FROM THEIR COMMITTED PROGENITOR
CC CELLS. IT ACTS AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT
CC MAY BE THE MAJOR PHYSIOLOGICAL REGULATOR OF CIRCULATING PLATELETS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DOMAIN: TWO-DOMAIN STRUCTURE WITH AN ERYTHROPOIETIN-LIKE N-
CC TERMINAL AND A SER/PRO/THR-RICH C-TERMINAL.
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D32207; BAA06906.1; -
DR InterPro: IPR001323; EPO_TPO.
DR InterPro: IPR003978; Thrombopoietin.
DR Pfam: PF00758; EPO_TPO; 1.
DR PRINTS: PR01485; THROMBOPTN.
DR PROSITE: PS00817; EPO_TPO; 1.
KM Cytokine; Glycoprotein; Hormone; Signal.
KW Cytokine; Glycoprotein; Hormone; Signal.
FT SIGNAL 1 21
FT CHAIN 22 326
FT DISULFID 28 172
FT CARBOHYD 50 106
FT CARBOHYD 197 197
FT CARBOHYD 206 206
FT CARBOHYD 235 235
FT CARBOHYD 249 249
FT CARBOHYD 256 256
SQ SEQUENCE 326 AA: 34556 MW: F99D7F77F8956FA2C CRC64;

Query Match 9.3%; Score 78; DB 1; Length 326;
Best Local Similarity 22.3%; Pred. No. 1.7;
Matches 37; Conservative 22; Mismatches 65; Indels 42; Gaps 6;

```

DB 122 LGALGLCTQLPPOG-----RTTAHKDPSALFLSLQOLKSKVR 161

RESULT 14

CH60_BABBA STANDARD: PRT: 543 AA.

AC P36357:

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE 60 kDa chaperonin (protein Cpn60) (groEL protein) (Immunoreactive protein Bb65) (Immunoreactive protein Bb63) (Heat shock protein 60) (HSP 60).

GN GROEL OR MODA OR GROEL OR 7B2 OR Bb65 OR Bb63.

OS Bartonella bacilliformis.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group; Bartonellaceae; Bartonella.

OX NCBI_TaxID-774;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-B13;

RA Schreiber M.G.;

RL Submitted (SEP-1992) to the EMBL/Genbank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Xu Y., Lu Z., Ihler G.;

RL Submitted (XXX-1993) to the EMBL/Genbank/DBJ databases.

RN [3]

RP SEQUENCE OF 1-36.

RX MEDLINE-91052092; PubMed-1700634;

RA Knobloch J., Schreiber M.G.;

RT "Bb65, a major immunoreactive protein of Bartonella bacilliformis," Am. J. Trop. Med. Hyg. 43:373-379(1990).

CC -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS CONDITIONS (BY SIMILARITY).

CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF 7 SUBUNITS (BY SIMILARITY).

CC -1- DISEASE: THIS PROTEIN IS A MAJOR ANTIGEN IN PATIENTS WITH BARTONELLOSIS, AN INFECTIOUS DISEASE ENDEMIC IN HIGH ALTITUDE VALLEYS OF THE ANDES.

CC -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

DR EMBL: Z15160; CAA78859.1; -

DR EMBL: M98257; A00146; -

DR PIR: A60146; A00146; -

DR PIR: S37039; S37039; -

DR HSPF: P06139; 1GRU.

DR InterPro: IPR001844; Chaperin_Cpn60.

DR InterPro: IPR002423; Cpn60_TCP1.1.

DR Pfam: PF00118; cpn60_TCP1.1.

DR PRINTS: PR00298; CHAPERONIN60.

DR PRINTS: PR00304; TCOMPLEXTCP1.

DR PROSITE: PS00296; CHAPERONINS_CPN60; 1.

DR Chaperone: ATP-binding; Antigen.

FT INIT_MET 0

FT CONFLICT 9 9 R -> N (IN REF. 3).

FT CONFLICT 15 15 L -> M (IN REF. 3).

FT CONFLICT 33 35 KGR -> VGV (IN REF. 3).

FT CONFLICT 240 240 A -> V (IN REF. 2).

FT CONFLICT 417 417 P -> A (IN REF. 2).

FT CONFLICT 485 485 R -> G (IN REF. 2).

SO SEQUENCE 543 AA; 57486 MM; ABB3407033A64D03 CMC64;

Query Match 9.1%; Score 76.5; DB 1; Length 543;

Best Local Similarity 27.0%; Pred. No. 4.6;

Matches 43; Conservative 24; Mismatches 59; Indels 33; Gaps 8;

DB 17 LEAKEAENTTTCACRCSL--NEN--ITVPTKVNFRYAKRNKNSXQAAVWVGCLA----- 68

DB 149 ISANGAEDIGKAIADMEKVGNEGVITVEAK-----TATLEVEVGKQFDRCY 198

QY 69 LSEAVLRGQALVNSQPEPLQLHVDKAVGSLRSLTLRALGAKREAISSPDAASA 128

DB 199 LSPYFVTNKKMMVDDDDY--ILIH-EKKLSNLSLPLVEAV-----AQSGK 244

QY 129 PLRTTADT-FKLRVYSNPLRGKLYTGACRTGDR 166

DB 245 PLLIAEDVEGEALATLVNKLKGLKIAVAKPGCDR 283

RESULT 15

TPO_MOUSE STANDARD: PRT: 356 AA.

AC P40226;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Thrombopoietin precursor (Megakaryocyte colony stimulating factor) (MPL) (Megakaryocyte leukemia virus oncogene ligand) (C-mpl ligand) (ML) (Megakaryocyte growth and development factor) (MGDF).

GN THPO.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID-10090;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC TISSUE-Kidney;

RX MEDLINE-94261207; PubMed-8202158;

RA Lok S., Kaushansky K., Holly R.D., Kuiper J.L., Lofton-Day C.E., Oort P.J., Grant F.J., Hapel M.D., Burkhart S.K., Kramer J.M., Bell L.A.N., Sprecher C.A., Blumberg H., Johnson R., Prunkard D., Ching A.F.T., Mathews S.L., Bailey M.C., Forstrom J.W., Buddle M.M., Osborn S.G., Evans S.J., Sheppard P.O., Presnell S.R., O'Hara P.J., Hagen F.S., Roth G.J., Foster D.C.,

RT "Cloning and expression of murine thrombopoietin cDNA and stimulation of platelet production in vivo." Nature 369:565-568(1994).

RT [2]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC TISSUE-Liver;

RX MEDLINE-94291201; PubMed-8020099;

RA Bartley T.D., Bogenberger J., Hunt P., Li Y.-S., Lu H.S., Martin F., Chang M.-S., Samal B.B., Nichol J.L., Swift S., Johnson M.J., Hsu R.-Y., Parker V.P., Suggs S., Skrine J.D., Merewether L.A., Closson C., Hsu E., Horkohl A., Choi E., Pangelinan M., Sun Y., Mar V., McNich J., Simonet L., Jacobsen F., Xie C., Shutter J., Chute H., Basu R., Selander L., Trollinger D., Siu L., Padilla D., Trail G., Elliott G., Izumi R., Covey T., Crouse J., Garcia A., Xu W., del Castillo J., Biron J., Cole S., Hu M.C.-T., Pacifici R., Ponting I., Sarris C., Wen D., Yung Y.P., Lin H., Rossman R.A.;

RT "Identification and cloning of a megakaryocyte growth and development factor that is a ligand for the cytokine receptor MPL." Cell 77:1117-1117(1994).

RL [3]

RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).

RX MEDLINE-95152076; PubMed-7849319;

RA Gurney A.L., Kuang W.-J., Xie M.-H., Malloy B.E., Eaton D.L., de Sauvage F.J.;

RT "Genomic structure, chromosomal localization, and conserved alternative splice forms of thrombopoietin." Blood 85:981-986(1995).

RL -1- FUNCTION: LINEAGE-SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION AND MATURATION OF MEGAKARYOCYTES FROM THEIR COMMITTED PROGENITOR CELLS. IT ACTS AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 02:35:14 ; Search time 23.1198 Seconds
(without alignments)
690.245 Million cell updates/sec

Title: US-09-813-775c-18

Perfect score: 843
Sequence: 1 APPRLICDSRVLELYLEAK.....NFLRGKLLKLTGECAGRTGDR 166

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	825	97.9	193	1 ZUHU	erythropoietin pre
2	743.5	88.2	192	1 JQ0173	erythropoietin pre
3	741.5	88.0	192	1 I84613	erythropoietin pre
4	695	82.4	188	1 I46083	erythropoietin pre
5	682	82.1	192	1 S28148	erythropoietin pre
6	675.5	80.1	195	2 JC7699	erythropoietin - r
7	672.5	79.8	194	1 I46401	erythropoietin pre
8	668	79.2	192	1 A24902	erythropoietin pre
9	665	78.9	190	1 I46578	erythropoietin - p
10	615	73.0	175	2 I46199	erythropoietin - d
11	94	11.2	353	2 G02729	thrombopoietin - h
12	93	11.0	353	2 I80105	thrombopoietin pre
13	90	10.7	286	2 A55530	megakaryocyte grow
14	87.5	10.4	346	2 AE0959	Solute binding rec
15	84	10.0	339	2 A83274	UDP-N-acetylpyruv
16	81	9.6	323	2 AB0323	ribonucleoside-dip
17	78	9.3	326	2 JC4125	thrombopoietin pre
18	77.5	9.2	242	2 AD1928	hypothetical prote
19	77.5	9.2	267	2 T08405	hypothetical prote
20	77	9.1	348	2 T35450	ABC transporter AT
21	77	9.1	451	2 S75569	hypothetical prote
22	77	9.1	480	2 S56639	ribosomal protein
23	76.5	9.1	544	2 S37039	groEL protein - Ba
24	76.5	9.1	1638	2 T30313	chemotaxis protein
25	76.5	9.1	2472	2 E83594	still frameshift p
26	76	9.0	356	2 S45330	thrombopoietin - m
27	75	8.9	552	2 S39765	chaperonin 60 - Co
28	75	8.9	263	2 B75361	WD-repeat family p
29	74.5	8.8	453	2 H97693	methylamine util112

30	74.5	8.8	455	2 AG2919	conserved hypothet
31	74	8.8	1089	2 S53978	PSEI protein - yea
32	73.5	8.7	897	2 A54696	EGF receptor subst
33	73	8.7	353	2 AG0138	guinolate synth
34	72.5	8.6	637	2 S75772	hypothetical prote
35	72.5	8.6	762	2 E87592	hypothetical prote
36	72	8.5	1846	2 T33079	hypothetical prote
37	71.5	8.5	220	2 E70749	probable hycp prot
38	71.5	8.5	296	2 A10443	probable 2-hydroxy
39	71.5	8.5	821	2 T49560	probable excision
40	71	8.4	243	2 AE1800	reductases homolog
41	71	8.4	300	2 E84310	hypothetical prote
42	71	8.4	354	2 B75355	hypothetical prote
43	70.5	8.4	547	2 G81787	glucose-6-phosphat
44	70.5	8.4	401	2 H83911	hypothetical prote
45	70.5	8.4	401	2 AF3341	precorrin-6y c5,15

ALIGNMENTS

RESULT 1

ZUHU erythropoietin precursor [validated] - human

C:Species: Homo sapiens (man)

C>Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 08-Dec-2000

C:Accession: A01855; A24744; A25384; A23210; S56178

R:Jacobs, K.; Shoemaker, C.; Ruderstorf, R.; Neill, S.D.; Kaufman, R.J.; Mulson, A.;

Nature 313, 806-810, 1985

A:Title: Isolation and characterization of genomic and cDNA clones of human erythropo

A:Reference number: A01855; M01D:851378959; PMID:3838366

A:Accession: A01855

A:Molecule type: mRNA; DNA

A:Residues: 1-193 <IAC>

A:Cross-references: GB:X02157; GB:X02158

R:Lin, F.K.; Suggs, S.; Lin, C.H.; Browne, J.K.; Smalling, R.; Eyrle, J.C.; Chen, K.K

Proc. Natl. Acad. Sci. U.S.A. 82, 7580-7584, 1985

A:Title: Cloning and expression of the human erythropoietin gene.

A:Reference number: A24744; M01D:86067948; PMID:3865178

A:Accession: A24744

A:Molecule type: DNA

A:Residues: 1-193 <LIN>

A:Cross-references: GB:M11319; M1D:9182197; PIDN:AA52400.1; PID:9182198

R:Li, P.H.; Everett, R.; Wang, F.F.; Arakawa, T.; Goldwasser, E.

J. Biol. Chem. 261, 3116-3121, 1986

A:Title: Structural characterization of human erythropoietin.

A:Reference number: A25384; M01D:86140080; PMID:3949763

A:Accession: A25384

A:Molecule type: protein

A:Residues: 28-86, 'Q', '87-193 <LAT>

A:Experimental source: urine

A:Note: forms with the carboxyl-terminal residue and the four carboxyl-terminal re

R:Vanagaw, S.; Hirade, K.; Ohnori, H.; Sasaki, R.; Chiba, H.; Ueda, M.; Goto, M.

J. Biol. Chem. 259, 2707-2710, 1984

A:Title: Isolation of human erythropoietin with monoclonal antibodies.

A:Reference number: A22210; M01D:84135751; PMID:6698989

A:Accession: A22210

A:Molecule type: protein

A:Residues: 28-29, 'X', '31-33, 'L', '35-50, 'X', '52-53, 'D', '55, 'G', '57 <YAN>

R:Matsumoto, S.; Ikura, K.; Ueda, M.; Sasaki, R.

Plant Mol. Biol. 27, 1163-1172, 1995

A:Title: Characterization of a human glycoprotein (erythropoietin) produced in cultur

A:Reference number: S56178; M01D:95284365; PMID:7766897

A:Accession: S56178

A:Molecule type: protein

A:Residues: 28-33, 'X', '35-37 <MTS>

C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live

C:Genetics:

A:Gene: GDB:EPO

A:Cross-references: GDB:119110; OMIM:133170

A:Map position: 7q21.3-7q22.1

A:Introns: 5/1; 53/3; 82/3; 142/3

C:Function:

A:Description: the primary inducer of erythrocyte formation
C:Superfamily: erythropoietin
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-193/Product: erythropoietin #status experimental <MAT>
F:34-187,56-60/Disulfide bonds: #status experimental
F:51,65,110/Binding site: carbohydrate (asn) (covalent) #status experimental
F:152/Binding site: carbohydrate (Ser) (covalent) #status experimental

Query Match 97.9%; Score 825; DB 1; Length 193;
Best Local Similarity 97.6%; Pred. No. 2,4e-75;
Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 APPRLICDSRYLERYLLAKEAENITTCGAHCISLNENITVPDTRKVNRYAMKRNKXSOQA 60
DB 28 APPRLICDSRYLERYLLAKEAENITTCGAHCISLNENITVPDTRKVNRYAMKRMVEGQQA 87

OY 61 VEWVWGLALLSEAVLRGALLVNSQPEPQLQHDKAVSGIRSLITLLRALGAQKEAIS 120
DB 88 VEWVWGLALLSEAVLRGALLVNSQPEPQLQHDKAVSGIRSLITLLRALGAQKEAIS 147

OY 121 PPDASAPLRTITADTRFKLFRVYSNLRGKLYTGECACRTGDR 166
DB 148 PPDASAPLRTITADTRFKLFRVYSNLRGKLYTGECACRTGDR 193

RESULT 2
JQ0173
erythropoietin precursor - crab-eating macaque
C:Species: Macaca fascicularis (crab-eating macaque)
C:Date: 07-Sep-1990 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999
C:Accession: JQ0173
R:Lin, F.K.; Lin, C.H.; Lai, P.H.; Browne, J.K.; Egrie, J.C.; Smalling, R.; Fox, G.M.;
Gene 44, 201-209, 1986
A:Title: Monkey erythropoietin gene: cloning, expression and comparison with the human
A:Reference number: JQ0173; MUID:87055236; PID:2877922
A:Accession: JQ0173
A:Molecule type: mRNA
A:Residues: 1-192 <LIN>
A:Cross-references: GB:M18189; GB:M15818; GB:M15819; GB:M18188; NID:g342093; PIDN:AAA368
A:Experimental source: kidney
C:Comment: This protein is the principal hormone involved in the regulation of erythrocy
C:Function: Erythropoietin is produced by kidney or liver of adult mammals and by liver
A:Description: the primary inducer of erythrocyte formation
C:Superfamily: erythropoietin
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-193/Product: erythropoietin #status experimental <MAT>
F:34-187,56-60/Disulfide bonds: #status predicted
F:51,65,110/Binding site: carbohydrate (asn) (covalent) #status predicted
F:152/Binding site: carbohydrate (Ser) (covalent) #status predicted

Query Match 88.2%; Score 743.5; DB 1; Length 192;
Best Local Similarity 89.2%; Pred. No. 3.6e-67;
Matches 148; Conservative 7; Mismatches 10; Indels 1; Gaps 1;

OY 1 APPRLICDSRYLERYLLAKEAENITTCGAHCISLNENITVPDTRKVNRYAMKRNKXSOQA 60
DB 28 APPRLICDSRYLERYLLAKEAENITTCGAHCISLNENITVPDTRKVNRYAMKRMVEGQQA 87

OY 61 VEWVWGLALLSEAVLRGALLVNSQPEPQLQHDKAVSGIRSLITLLRALGAQKEAIS 120
DB 88 VEWVWGLALLSEAVLRGALLVNSQPEPQLQHDKAVSGIRSLITLLRALGAQKEAIS 146

OY 121 PPDASAPLRTITADTRFKLFRVYSNLRGKLYTGECACRTGDR 166
DB 147 LPDASAPLRTITADTRFKLFRVYSNLRGKLYTGECACRTGDR 192

RESULT 3
184613
erythropoietin precursor - rhesus macaque

C:Species: Macaca mulatta (rhesus macaque)
C:Date: 02-Aug-1996 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999
C:Accession: 184613
R:Men, D.; Boissel, J.
Blood 82, 1507-1516, 1993
A:Title: Erythropoietin structure-function relationships: High degree of sequence hom
A:Reference number: 146083; MUID:93372347; PMID:8364201
A:Accession: 184613
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-192 <RES>
A:Cross-references: GB:L10609; NID:g342095; PIDN:AAA36842.1; PID:g342096
C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live
C:Function:
A:Description: the primary inducer of erythrocyte formation
C:Superfamily: erythropoietin
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-193/Product: erythropoietin #status predicted <MAT>
F:34-187,56-60/Disulfide bonds: #status predicted
F:51,65,110/Binding site: carbohydrate (asn) (covalent) #status predicted
F:152/Binding site: carbohydrate (Ser) (covalent) #status predicted

Query Match 88.0%; Score 741.5; DB 1; Length 192;
Best Local Similarity 88.6%; Pred. No. 5.8e-67;
Matches 147; Conservative 8; Mismatches 10; Indels 1; Gaps 1;

OY 1 APPRLICDSRYLERYLLAKEAENITTCGAHCISLNENITVPDTRKVNRYAMKRNKXSOQA 60
DB 28 APPRLICDSRYLERYLLAKEAENITTCGAHCISLNENITVPDTRKVNRYAMKRMVEGQQA 87

OY 61 VEWVWGLALLSEAVLRGALLVNSQPEPQLQHDKAVSGIRSLITLLRALGAQKEAIS 120
DB 88 VEWVWGLALLSEAVLRGALLVNSQPEPQLQHDKAVSGIRSLITLLRALGAQKEAIS 146

OY 121 PPDASAPLRTITADTRFKLFRVYSNLRGKLYTGECACRTGDR 166
DB 147 LPDASAPLRTITADTRFKLFRVYSNLRGKLYTGECACRTGDR 192

RESULT 4
146083
erythropoietin precursor - cat (fragment)
C:Species: Felis silvestris catus (domestic cat)
C:Date: 16-Aug-1996 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999
C:Accession: 146083
R:Men, D.; Boissel, J.
Blood 82, 1507-1516, 1993
A:Title: Erythropoietin structure-function relationships: High degree of sequence hom
A:Reference number: 146083; MUID:93372347; PMID:8364201
A:Accession: 146083
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-188 <MEN>
A:Cross-references: GB:L10606; NID:g163820; PIDN:AAA30807.1; PID:g163821
C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live
C:Function:
A:Description: the primary inducer of erythrocyte formation
C:Superfamily: erythropoietin
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F:1-27/Domain: signal sequence (fragment) #status predicted <SIG>
F:23-186/Product: erythropoietin #status predicted <MAT>
F:29-183,51-55/Disulfide bonds: #status predicted
F:46,60,105/Binding site: carbohydrate (asn) (covalent) #status predicted
F:148/Binding site: carbohydrate (Ser) (covalent) #status predicted

Query Match 82.4%; Score 695; DB 1; Length 188;
Best Local Similarity 82.5%; Pred. No. 2.6e-62;
Matches 137; Conservative 8; Mismatches 21; Indels 0; Gaps 0;

OY 1 APPRLICDSRYLERYLLAKEAENITTCGAHCISLNENITVPDTRKVNRYAMKRNKXSOQA 60
DB 23 APPRLICDSRYLERYLLAKEAENITTCGAHCISLNENITVPDTRKVNRYAMKRMVEGQQA 82

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 8, 2003, 05:18:00 ; Search time 49.4596 Seconds
(without alignments)
1717.360 Million cell updates/sec

Title: US-09-813-775c-2

Perfect score: 998
Sequence: 1 MGVECPAMWLLSLSLP.....NFLRKKLYTGACRTGDR 193

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 389086 seqs, 220051671 residues

Total number of hits satisfying chosen parameters: 778172

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame_plus.model -DEV=xlh
-O=cgcn2.1/USPTO.spool/US09813775/runat.07012003.153123.23862/app.query.fasta_1.718
-DB=Published Applications NA -GPM=fastlap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blousum62
-TRANS=human40.cdli -LIST=45 -DOCLIGN=200 -THR=SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pico -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09813775.@CGN_1.1.25 @runat.07012003.153123.23862
-NCPU=6 -ICPU=3 -NO_XLPEXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_NA:
1: /cgcn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgcn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgcn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgcn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgcn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgcn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgcn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgcn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgcn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgcn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgcn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgcn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgcn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
14: /cgcn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	709	71.0	2426	10	US-09-975-063-1
2	709	71.0	2426	12	US-10-011-858-1
3	327	32.8	423	10	US-09-864-761-16087
4	286	28.7	191	10	US-09-864-761-32584

5	90	9.0	3149	10	US-09-729-674-127	Sequence 127, App
6	84	8.4	2124	10	US-09-731-872-39	Sequence 39, Appl
7	84	8.4	2300	9	US-09-992-598-137	Sequence 137, Appl
8	84	8.4	2300	9	US-09-989-293A-137	Sequence 137, App
9	84	8.4	2300	9	US-09-989-735-137	Sequence 137, App
10	84	8.4	2300	9	US-09-990-444-137	Sequence 137, App
11	84	8.4	2300	9	US-09-989-730-137	Sequence 137, App
12	84	8.4	2300	9	US-09-990-436-137	Sequence 137, App
13	84	8.4	2300	9	US-09-991-181-137	Sequence 137, App
14	84	8.4	2300	9	US-09-993-687-137	Sequence 137, App
15	84	8.4	2300	9	US-09-989-734-137	Sequence 137, App
16	84	8.4	2300	10	US-09-989-723-137	Sequence 137, App
17	84	8.4	2300	10	US-09-989-723-137	Sequence 137, App
18	84	8.4	2300	10	US-09-989-279-137	Sequence 137, App
19	84	8.4	2300	10	US-09-989-727-137	Sequence 137, App
20	84	8.4	2300	10	US-09-989-731-137	Sequence 137, App
21	84	8.4	2300	10	US-09-989-732-137	Sequence 137, App
22	84	8.4	2300	10	US-09-991-073-137	Sequence 137, App
23	84	8.4	2300	10	US-09-990-442-137	Sequence 137, App
24	84	8.4	2300	10	US-09-991-163-137	Sequence 137, App
25	84	8.4	2300	10	US-09-993-604-137	Sequence 137, App
26	84	8.4	2300	10	US-09-990-456-137	Sequence 137, App
27	84	8.4	2300	10	US-09-989-721-137	Sequence 137, App
28	84	8.4	2300	12	US-10-052-586-115	Sequence 115, App
29	84	8.4	2357	9	US-10-098-841-62	Sequence 62, Appl
30	83.5	8.4	1786	10	US-09-969-708-492	Sequence 492, App
31	80	8.0	4308	9	US-09-978-295A-527	Sequence 527, App
32	80	8.0	4308	9	US-09-978-697-527	Sequence 527, App
33	80	8.0	4308	9	US-09-978-192A-527	Sequence 527, App
34	80	8.0	4308	9	US-09-999-832A-527	Sequence 527, App
35	80	8.0	4308	9	US-09-978-189-527	Sequence 527, App
36	79.5	8.0	2863	9	US-10-136-224-2	Sequence 2, Appl1
37	79.5	8.0	3686	9	US-10-136-224-3	Sequence 3, Appl1
38	79.5	8.0	3885	10	US-09-815-224-3	Sequence 9626, Ap
39	79.5	8.0	3940	9	US-10-136-224-1	Sequence 1, Appl1
40	77.5	7.8	1372	10	US-09-925-301-575	Sequence 575, App
41	77.5	7.8	2442	9	US-09-964-899-24	Sequence 24, Appl
42	77.5	7.8	78056	9	US-10-109-551-1	Sequence 1, Appl1
43	76.5	7.7	1663	10	US-09-070-927A-277	Sequence 277, App
44	76.5	7.7	3090	10	US-09-768-436-1	Sequence 1, Appl1
45	76.5	7.7	4006	10	US-09-768-436-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-09-975-063-1
Sequence 1, Application US/0975063
Patent No. US20020045255A1
GENERAL INFORMATION:
APPLICANT: POWELL, Jerry S.
TITLE OF INVENTION: HUMAN ERYTHROPOIETIN GENE: HIGH LEVEL
EXPRESSION IN STABLY TRANSFECTED MAMMALIAN CELLS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dorsey & Whitney, LLP
STREET: Suite 3400, 1420 Fifth Avenue, U.S. Bank Centre
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6/7/8
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/975,063
FILING DATE: 10-Oct-2001
ATTORNEY/AGENT INFORMATION:
NAME: Roberts, Mark W.
REGISTRATION NUMBER: 46,160
REFERENCE/DOCKET NUMBER: 500582.03 (112893.109)

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 206-903-8728
 TELEFAX: 206-903-8820
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2426 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-975-063-1

Alignment Scores:
 Pred. No.: 2,93e-78 Length: 2426
 Score: 709.00 Matches: 186
 Percent Similarity: 35.80% Conservat: 3
 Best Local Similarity: 35.23% Mismatches: 4
 Query Match: 71.04% Indels: 337
 Gaps: 3

US-09-813-775C-2 (1-193) x US-09-975-063-1 (1-2426)

```

OY 1 MetGlyValHisGluCysProAlaTrpLeuTrpLeuLeuSerLeuLeuSerLeuPro 20
    ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 625 CTATCTGTCTTAAGAAATGCTCTGCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCT 684
OY 21 LeuGlyLeuProValLeuGlyAlaProProAlaGlyLeuIleCysAspSerArgValLeuGlu 40
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 685 CTGGGCGCTCCAGCTCGCTGGGCGCCGCCACACGCTCATCTGTGACAGCGGAGCTCGCAG 744
OY 41 ArgTrpLeuLeuGluAlaGlyGluAlaGluAlaGluAlaGluAlaGluAlaGluAla 53
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 745 AGGTACCTCTTGGAGGCCAAGAGGCCGAGAAATATCATCGGTAGAGACCCCTTCCAGCAG 804
OY 53 :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::: 53
Db 805 ATTCCAGAGAACTACACGCTCAGGCGCTTCAGGAGACTCTCCAGATCCAGAACTGCGCA 864
OY 53 :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::: 53
Db 865 CTTCGTTGGGGTGGAGTTGGAGAGCTAGACACTGCCGCCCTACATATAGATAGCTGCG 924
OY 53 :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::: 53
Db 925 TGGCCCAAAACCATACCTGGAAACTAGGCAAGAGCAAAAGCAGAGATCTAGCGCTGT 984
OY 54 :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::: 984
Db 985 GGGCAGGGCGCAGAGCTTCAGGGACCTTGACTCCCGGGGCTGTGTCATTTTCAGAGCGGG 1044
OY 55 yCysAlaGluHisCysSerLeuAsnGluAsnIleThrValProAspThrLysValAsnPh 75
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1045 CTGTCTCAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1104
OY 75 eTryAlaTrpLysArgMetGluValArg:::::::::::::::::::::::::::: 84
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1105 CTATGCTGGAGAGAGATGGAGGTAG- TTCTTTTTTTTTTTTTTTTTTTCTTTTGA 1163
OY 84 :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::: 84
Db 1164 GAATCTCATTTTGGAGCTGATTTTGATGAAGAAGGAGAAATGATGAGGAAAGTAATA 1223
OY 84 :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::: 84
Db 1224 TGGAGCAGCAGATGAGGCTGCTGGGCGCAGAGGCTCAGCTATATATCCAGAGCTGA 1283
OY 84 :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::: 84
Db 1284 GATGCGCAGATGGAGAAATTTGCTTGAGCCCGGAGTTTCAGACCAACCTAGCAGACATA 1343
OY 84 :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::: 84

```

```

Db 1344 GTGAGTCCCGCATCTCTCAAAACATTTAAAAAATTACTAGGTGAGTGTGATGCT 1403
OY 84 :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::: 84
Db 1404 GGTAGTCCAGATATTTTGAAGGCTGAGGCGGAGAGATCGTGGAGCCAGGAATTTGAG 1463
OY 84 :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::: 84
Db 1464 GGTGAGTGAAGTGTGATCATCACACACTGAATCCAGCTCAGTACAGAGTGAAGCCCTG 1523
OY 84 :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::: 84
Db 1524 TCTCAAAAAAGAAAAAGAAAAAATAATGAGGCTGTATGAAATACGTTCAATTAT 1583
OY 84 :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::: 84
Db 1584 CATTCAGTCACTCACTCACTCACTATTCATTTCATTTCATTCAAGAGTTCATTGATAC 1643
OY 84 :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::: 84
Db 1644 CTTCGTGTTGCTCAGCTGTGCTGCTGGGCTGCTGAGGGGAGAGAGGAGTGCAT 1703
OY 85 :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::: 85
Db 1704 CCTCAGCTGACTCCAGAGTCCACTCCCTGATAGTGGCGCAGCAGCGCTGAGAGTCTG 1763
OY 91 pGlnGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeuValAsnSe 111
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1764 GAGAGGCGTGGCGCTGCTGCTGCGAAGCTGCTCGGGGGCCAGGCGCTTGTGTAGACAC 1823
OY 111 rSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSerGlyLeuArgSe 131
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1824 TTCCACGCGGTGGAGCCCTCAGCTGATGTGATTAAGCCGTCAGTGGCTTCGCGAG 1883
OY 131 rLeuThrTrpLeuLeuArgAlaLeuGlyAla:::::::::::::::::::: 141
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1884 CCTCACACTCTGCTTGGGCTGTGGAGC-CCAGGTGATGAGAGCGACACTTCTGCT 1942
OY 141 :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::: 141
Db 1943 TGCCCTTTCTGAAGAGGAGAAAGGTCTGTGAAGATACAGAGACTGCTGATT 2002
OY 142 :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::: 142
Db 2003 CCTTCCTTTCTGAGCACTGAGCGACCTCTGTTCTCTCTTGGCAGAGAAAGCCAT 2062
OY 146 eSerProProAlaAlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPhear 166
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 2063 CTCCCTCCAGATGGCGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2122
OY 166 gLysLeuPhearGlyValTyrSerAsnPhLeuArgGlyLysLeuLysLeuTyrThrGly 186
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 2123 CAACCTCTCCAGTCTACTCCAAATTTCTCCGGGAAAGCTGAAGCTGTACAGAGGGA 2182
OY 186 uAlaCysArgThrGlyAspArg 193
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 2183 GGCCTGACAGACAGGGGACAGA 2204
    |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

RESULT 2
 US-10-011-858-1

; Sequence 1, Application US/10011858
 ; Patent No. US20020137145A1

GENERAL INFORMATION:

APPLICANT: POWELL, Jerry S.

TITLE OF INVENTION: HUMAN ERYTHROPOIETIN GENE: HIGH LEVEL

EXPRESSION IN STABLY TRANSFECTED MAMMALIAN CELLS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dorsey & Whitney, LLP

STREET: Suite 3400, 1420 Fifth Avenue, U. S. Bank Centre

CITY: Seattle

STATE: Washington

COUNTRY: USA

```

1      ZIP: 98101
2
3      COMPUTER READABLE FORM:
4
5      MEDIUM TYPE: Floppy disk
6
7      COMPUTER: IBM PC compatible
8
9      OPERATING SYSTEM: PC-DOS/MS-DOS
10     SOFTWARE: MS-WORD
11
12     CURRENT APPLICATION DATA:
13
14     APPLICATION NUMBER: US/10/011,858
15
16     FILING DATE: 05-NO. US20020137145A1-2001
17
18     ATTORNEY/AGENT INFORMATION:
19
20     NAME: Roberts, Mark W.
21
22     REGISTRATION NUMBER: 46,160
23
24     REFERENCE/DOCKET NUMBER: 500582.13
25
26     TELECOMMUNICATION INFORMATION:
27
28     TELEPHONE: 206-903-8728
29
30     TELEFAX: 206-903-8820
31
32     INFORMATION FOR SEQ ID NO: 1:
33
34     SEQUENCE CHARACTERISTICS:
35
36     LENGTH: 2426 base pairs
37
38     TYPE: nucleic acid
39
40     STRANDEDNESS: single
41
42     TOPOLOGY: linear
43
44     MOLECULE TYPE: DNA (genomic)
45
46     SEQUENCE DESCRIPTION: SEQ ID NO: 1:
47
48     US-10-011-858-1

```

Sequence 16087, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Penn, Sharon G.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmics-X-1
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 16087
LENGTH: 423
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AF053356.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.93
US-09-864-761-16087

Alignment Scores:
Pred. No.: 6,638-32 Length: 423
Score: 327.00 Matches: 78
Percent Similarity: 62.40% Conservative: 0
Best Local Similarity: 62.40% Mismatches: 2
Query Match: 32.77% Indels: 46
DB: 10 Gaps: 1

US-09-813-775c-2 (1-193) x US-09-864-761-16087 (1-423)

QY 83 VALARGINGLALVALGLUVALTRPGINGLYLEUALALEULEUSERG1UALAVALLEN 102
DB 376 GTGGGAGACGAGCCGTAAAGTCTGGCAGGCGCTGGCCCTCTGTGCGAAGCTGTCTG 317

QY 103 ARGGLYLALALEULEVALANSERSEGINPROTRPGLUPLROLEUNG1LNUHISVAL 122
DB 316 CGGGCCAGAGCCCTGTGGTCAACTCTCCAGCGGTGGAGCCCTGCAGCTGATGTG 257
QY 123 ASPLYLALVALSERGLYLEUARGLSERLEUTHNTHLEULNARGALALEULYALA---- 141
DB 256 GATAAAGCGTAGTGGCTTCGACGCCACACCTCTCTGGGCTGTGGAGC-CCA 198
QY 141 ----- 141
DB 197 GGTGATGAGAGCGACACTTCTGCTGGCTTTCTGTAGAAGGAGGAGAGGCTTTCG 138
QY 141 ----- 141
DB 137 TAAGAGTACAGGAAGTCTGCTATTCCTTCTGTGACACTGACGACCTCTG 78
QY 142 -----G1ULYSLUALALLESERPROPRORASPLALASERLALALPROLEU 157
DB 77 TTTTCTCTTGGCAGAGAGCCATCTCCCTCAGATGGGCTCAGCTCTCCACTC 18
QY 158 ARGTHRIETHRALA 162
DB 17 CGAACAATCACTGCT 3

RESULT 4
US-09-864-761-32584/C
Sequence 32584, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aecmics-X-1
CURRENT FILING DATE: US/09/864,761
PRIOR APPLICATION NUMBER: 2001-05-23
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203

; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 32584
 ; LENGTH: 191
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AF053356.1
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.93
 ; OTHER INFORMATION: NT HIT: AF202314.1, EVALUATE 1.00e-104
 ; OTHER INFORMATION: EST_HUMAN HIT: AA662379.1, EVALUATE 4.00e-57
 ; OTHER INFORMATION: SWISSPROT HIT: P01588, EVALUATE 1.00e-27
 ; US-09-864-761-32584

Alignment Scores:
 Pred. No.: 2,51e-27 Length: 191
 Score: 286.00 Matches: 59
 Percent Similarity: 98.33% Conservative: 0
 Best Local Similarity: 98.33% Mismatches: 1
 Query Match: 28.66% Indels: 0
 Gaps: 0

US-09-813-775C-2 (1-193) x US-09-864-761-32584 (1-191)

QY 83 ValArgGlnGlnAlaValAlaGluValTrpGlnGlyLeuAlaLeuSerGluAlaValLeu 102
 Db 180 GTGGGAGAGAGCGCTAGAACTGTGGCAGGCGCTGGCTGTGTGGAAGCTGTCTGT 121
 QY 103 ArgGlyGlnAlaLeuLeuValAsnSerSerGlnProTrpGlnProLeuGlnLeuHisVal 122
 Db 120 CGGGGCCAGGCGCTGTGTGCTCAACTGTCCACCGCGAGCGCCGCGACGTCATGTG 61
 QY 123 AspLysAlaValAlaSerGlyLeuArgSerLeuThrLeuLeuArgAlaLeuGlyAlaGln 142
 Db 60 GATAAAGCGCTCACTGCGCCCTTCGACGCTCACCACTGTCTGTGGGCTCTGGAGCCAG 1

RESULT 5
US-09-729-674-127

; Sequence 127, Application US/09729674
 ; Patent No. US20010039335A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jacobs, Kenneth
 ; APPLICANT: McCoy, John M.
 ; APPLICANT: Lavallee, Edward R.
 ; APPLICANT: Collins-Racle, Lisa A.
 ; APPLICANT: Evans, Cheryl
 ; APPLICANT: Merberg, David
 ; APPLICANT: Treacy, Maurice
 ; APPLICANT: Agostino, Michael J.
 ; APPLICANT: Steindinger II, Robert J.
 ; APPLICANT: Spaulding, Vikki
 ; APPLICANT: Wong, Gordon G.
 ; APPLICANT: Clark, Hilary
 ; APPLICANT: Fechtel, Kim
 ; APPLICANT: Genetics Institute, Inc.
 ; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
 ; FILE REFERENCE: 6055-64X
 ; CURRENT APPLICATION NUMBER: US/09/729,674
 ; CURRENT FILING DATE: 2000-12-04
 ; PRIOR APPLICATION NUMBER: 09/539,330
 ; PRIOR FILING DATE: 2000-03-30
 ; NUMBER OF SEQ ID NOS: 283
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 127
 ; LENGTH: 3149
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-729-674-127

Alignment Scores:
 Pred. No.: 0.397 Length: 3149
 Score: 90.00 Matches: 48

Percent Similarity: 44.03% Conservative: 22
 Best Local Similarity: 30.19% Mismatches: 63
 Query Match: 9.02% Indels: 26
 Gaps: 9

US-09-813-775C-2 (1-193) x US-09-729-674-127 (1-3149)

QY 6 CysProAlaTrpLeuTrpLeuLeuLeuSerLeuSerLeu-----ProLeuGlyLeu 23
 Db 364 TGCCCT-----CTGCTACTTCTTCAATAGGCTCCCGCAAGGCGCTGT 405
 QY 24 ProValLeuGlyAlaProProAlaGlyLeuIleCysAspSerArgValLeuGlyAlaGlyTyrLeu 43
 Db 406 TGTCTCTTGGTTCGCTCCGCCCAAGCTAATCTGTGAGAAACACATCAACAGAACCTTA 465
 QY 44 LeuGluAlaLysGluAlaGluAsnIleThr-----ThrGlyCysAlaGluHisCys 60
 Db 466 TTGGATATTTTGGCAGATGGCACCATTGTTCAAAGTTGAGTGGAGTGCATGCAAGATGCC 525
 QY 61 Ser-----LeuAsnGluAsnIleThrValProAspThrLysValAsnPheTyrAlaTrp 78
 Db 526 AGCAAGCTTGTGCAGATATATGGCTGTGTGTAGGGGGTGCCTGGACCTC----- 576
 QY 79 LysArgMetGluValArgGlnGlnAlaValAlaGluValTrpGlnGlyLeuAlaLeu----- 96
 Db 577 CGATACCTTACGACGAGCGGAGAAACAAATTCCTGTGTAATGGCGCTTACGCTGAAGTCC 636
 QY 97 LeuSerGluAlaValLeuValG-----GlyGlnAlaLeuValAsnSerSerGln 113
 Db 637 CTGCTGAGACTGTGTTCGAACCTTCCCTGACACAGTCCCTTCACTGCTGTGCGTGGCAGCAAC 696
 QY 114 ProTrpGlnProLeuGlnLeuHisValAspLysAlaValSerGlyLeuValG----- 130
 Db 697 ---TGGATGGTGGAGCTCCACAGAGACCAAGTAAATTATGCTGGCAGGATGCCAG 753
 QY 131 ---SerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGluAlaIleSerPro 148
 Db 754 ATTTCAGTGGCTCTTCTTTCATCTTCTGATACCTTCTGATACCTTCTCTAGAAATTCACCT 810

RESULT 6

; US-09-731-872-39
 ; Sequence 39, Application US/09731872
 ; Patent No. US20020102604A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dumas Milne Edwards, Jean Baptiste
 ; APPLICANT: Bouquelere, Lydie
 ; APPLICANT: Jobert, Severin
 ; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
 ; FILE REFERENCE: 78.US3.REG
 ; CURRENT APPLICATION NUMBER: US/09/731,872
 ; CURRENT FILING DATE: 2000-12-07
 ; PRIOR APPLICATION NUMBER: US 60/169,629
 ; PRIOR FILING DATE: 1999-12-08
 ; PRIOR APPLICATION NUMBER: US 60/187,470
 ; PRIOR FILING DATE: 2000-03-06
 ; NUMBER OF SEQ ID NOS: 482
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 39
 ; LENGTH: 2124
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 505..1590
 ; NAME/KEY: sig_peptide
 ; LOCATION: 505..624
 ; OTHER INFORMATION: Von Heijne matrix
 ; OTHER INFORMATION: score 8.5056444915604
 ; OTHER INFORMATION: seq VVMILMLTLVLG/MV
 ; US-09-731-872-39

Alignment Scores:
 Pred. No.: 1.22 Length: 2124

Oy 118 uGlnLeuHISValAspIysAlaValSerGlyLeuArgSerLeuThrThrLeuLeuArgAl 138
Db 1277 -----AGCTGCTCATGCATGAG----- 1294
Oy 138 aleuGlyAlaGlnIysGluAlaIleSerProProAspAlaIaSerAlaAlaProLeuAr 158
Db 1295 -CTGCATGCCCCGAGCATGCAGCGTACCTCTTAGGCCAGGTCTCTCCCAAGCTG 1353
Oy 158 gThrIleThrAlaAspThrPheArgIysLeuPheArgValItyrSerAsnPhe----- 175
Db 1354 GGCTCCTTGGTGGCGCTCA-----TTCAAGTGTACTACTCTTTTACCTAANG 1401
Oy 176 -----leuArgGlyIysLeuIysLeuTyThrGlyGluAlaCys 188
Db 1402 GTGTCTCATGTTGTGGCTTCTATAGCTCTCCACTCTTCCGAGCGCTGC 1450

RESULT 8
US-09-989-293A-137
Sequence 137, Application US/09989293A
Patent No. US20020177164A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, V. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tamas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PLC66
CURRENT APPLICATION NUMBER: US/09/989,293A
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907

APPLICANT: Kljavin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P27301C61
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738

PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24

1	PRIOR APPLICATION NUMBER: 60/089801
2	PRIOR FILING DATE: 1998-06-18
3	PRIOR APPLICATION NUMBER: 60/089907
4	PRIOR FILING DATE: 1998-06-18
5	PRIOR APPLICATION NUMBER: 60/089908
6	PRIOR FILING DATE: 1998-06-18
7	PRIOR APPLICATION NUMBER: 60/089947
8	PRIOR FILING DATE: 1998-06-19
9	PRIOR APPLICATION NUMBER: 60/089948
10	PRIOR FILING DATE: 1998-06-19
11	PRIOR APPLICATION NUMBER: 60/089952
12	PRIOR FILING DATE: 1998-06-22
13	PRIOR APPLICATION NUMBER: 60/090254
14	PRIOR FILING DATE: 1998-06-22
15	PRIOR APPLICATION NUMBER: 60/090349
16	PRIOR FILING DATE: 1998-06-23
17	PRIOR APPLICATION NUMBER: 60/090355
18	PRIOR FILING DATE: 1998-06-23
19	PRIOR APPLICATION NUMBER: 60/090429
20	PRIOR FILING DATE: 1998-06-24
21	PRIOR APPLICATION NUMBER: 60/090431
22	PRIOR FILING DATE: 1998-06-24
23	PRIOR APPLICATION NUMBER: 60/090435
24	PRIOR FILING DATE: 1998-06-24
25	PRIOR APPLICATION NUMBER: 60/090444
26	PRIOR FILING DATE: 1998-06-24
27	PRIOR APPLICATION NUMBER: 60/090445
28	PRIOR FILING DATE: 1998-06-24
29	PRIOR APPLICATION NUMBER: 60/090472
30	PRIOR FILING DATE: 1998-06-24
31	PRIOR APPLICATION NUMBER: 60/090535
32	PRIOR FILING DATE: 1998-06-24
33	PRIOR APPLICATION NUMBER: 60/090540
34	PRIOR FILING DATE: 1998-06-24
35	PRIOR APPLICATION NUMBER: 60/090542
36	PRIOR FILING DATE: 1998-06-24
37	PRIOR APPLICATION NUMBER: 60/090557
38	PRIOR FILING DATE: 1998-06-24
39	PRIOR APPLICATION NUMBER: 60/090676
40	PRIOR FILING DATE: 1998-06-25
41	PRIOR APPLICATION NUMBER: 60/090678
42	PRIOR FILING DATE: 1998-06-25
43	PRIOR APPLICATION NUMBER: 60/090678
44	PRIOR FILING DATE: 1998-06-25
45	PRIOR APPLICATION NUMBER: 60/090690
46	PRIOR FILING DATE: 1998-06-25
47	PRIOR APPLICATION NUMBER: 60/090694
48	PRIOR FILING DATE: 1998-06-25
49	PRIOR APPLICATION NUMBER: 60/090655
50	PRIOR FILING DATE: 1998-06-25
51	PRIOR APPLICATION NUMBER: 60/090656
52	PRIOR FILING DATE: 1998-06-25
53	PRIOR APPLICATION NUMBER: 60/090656
54	PRIOR FILING DATE: 1998-06-25
55	PRIOR APPLICATION NUMBER: 60/090862
56	PRIOR FILING DATE: 1998-06-26
57	PRIOR APPLICATION NUMBER: 60/090863
58	PRIOR FILING DATE: 1998-06-26
59	PRIOR APPLICATION NUMBER: 60/091360
60	PRIOR FILING DATE: 1998-07-01
61	PRIOR APPLICATION NUMBER: 60/091478
62	PRIOR FILING DATE: 1998-07-02
63	PRIOR APPLICATION NUMBER: 60/091544
64	PRIOR FILING DATE: 1998-07-01
65	PRIOR APPLICATION NUMBER: 60/091519
66	PRIOR FILING DATE: 1998-07-02
67	PRIOR APPLICATION NUMBER: 60/091626
68	PRIOR FILING DATE: 1998-07-02
69	PRIOR APPLICATION NUMBER: 60/091633
70	PRIOR FILING DATE: 1998-07-02
71	PRIOR APPLICATION NUMBER: 60/091978
72	PRIOR FILING DATE: 1998-07-07
73	PRIOR APPLICATION NUMBER: 60/091982

Alignment Scores:	
Pred. No.:	1.38
Score:	84.00
Percent Similarity:	38.25%
Best Local Similarity:	22.58%
Query Match:	8.42%
DB:	9
	Gaps:
	10
	Indels:
	63
	Mismatches:
	71
	Conservative:
	34
	Matches:
	49
	Length:
	2300

OY	3	ValHisGluCysProAlaTrpLeuTrpLeuLeuSerLeuLeuSerLeuProLeuGly	22
		: : : : : :	
Db	898	CTGACACATGATGATCTCTTCCTGGGGTTCTGCTCTCTGTTGATGACCTAGGT	957
OY	23	LeuPro----ValLeuGluAlaProProArgLeuLeuCysAspSerArgValLeuGluArg	41
		: : : : : : :	
Db	958	CTGGCCCGCATGTTCTCCGCTACCTGGGAGAGCTGATGACACCCCGCTGTCGACAC	1017
OY	42	TyrLeuLeuGluAlaLysGluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSer	61
		: : : : : : :	
Db	1018	---CTGGAG-----GAGCAGCTCTACTGCTGACGCCCTTGGAGGAGGCACC	1055
OY	62	LeuAsnGluAsnIleThrValProAspThrLysValAsnPheTyrAlaTrpLysArgMet	81
		: : : : : : :	
Db	1060	CTGACCCCGCAGGATCTGTAATCCTACTTCTCC-----TGCTGGCTGCTTTA	1104
OY	82	GluVal-----	83
		: : : : : : : :	
Db	1105	GACATGGAGCTGCTACACACAGACAGCTCTGCGCTGTGCACACACAGAGGCTCTGCTGAG	1164
OY	84	ArgGlnGlnAlaValGluValTrpGln-----GlyLeuAlaLeuLeuSer	98

Db 1165 AAGAGGGGAAAGGCTTACGCTTGGCAACGGAACCTGGGCTACCCCTGGCTATGCTGTGC 1224
Qy 99 GluaIaValIeuArGlyGlyGlnAlaIeuLeuValaSer-SerGlnProTPGluProLe 118
Db 1225 TTGGTGGGCGAGGGGCTCTGTGTCTATTTGGGCATCCACATCTCTG----- 1276
Qy 118 uGlnLeuHisValaSplysAlaValSerGlyLeuArGlySerLeuThrLeuLeuArGAl 138
Db 1277 -----ACGCTGCTCATCGATGAGC----- 1294
Qy 138 aleuGlyAlaGlnLysGluAlaIleSerProProAspAlaIaSerAlaIaProLeuAr 158
Db 1295 -CTGCCATGCCCGAGGATCAGAGTACTCTAGGCGAGTCTCTTCCAAAGCTG 1353
Qy 158 gThrIleThrAlaAspThrPheArGlyLeuPheArGlyValTyrSerAsnPe----- 175
Db 1354 GGCCTCTTGGTGGCGTCA-----TTGAGGTGTACATCATTTTACTTAATG 1401
Qy 176 -----LeuArGlyLysLeuLysLeuTyrThrGlyAlaLys 188
Db 1402 GTGTCTCAGTGTGTGGCTTCTATAGCTCTCAGCTTCCGAGCCTGC 1450
RESULT 12
US-09-990-436-137
; Sequence 137, Application US/09990436
; Publication No. US20020198148A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Bolstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Matanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2730P1C14
; CURRENT APPLICATION NUMBER: US/09/990,436
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28

; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600

APPLICANT: Gerber, Hanspeter
APPLICANT: Gertitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PIC53
CURRENT APPLICATION NUMBER: US/09/991,181
CURRENT FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24


```

      / PRIOR FILING DATE: 1998-06-24
      / PRIOR APPLICATION NUMBER: 60/090445
      / PRIOR FILING DATE: 1998-06-24
      / PRIOR APPLICATION NUMBER: 60/090472
      / PRIOR FILING DATE: 1998-06-24
      / PRIOR APPLICATION NUMBER: 60/090535
      / PRIOR FILING DATE: 1998-06-24
      / PRIOR APPLICATION NUMBER: 60/090540
      / PRIOR FILING DATE: 1998-06-24
      / PRIOR APPLICATION NUMBER: 60/090542
      / PRIOR FILING DATE: 1998-06-24
      / PRIOR APPLICATION NUMBER: 60/090557
      / PRIOR FILING DATE: 1998-06-24
      / PRIOR APPLICATION NUMBER: 60/090676
      / PRIOR FILING DATE: 1998-06-25
      / PRIOR APPLICATION NUMBER: 60/090678
      / PRIOR FILING DATE: 1998-06-25
      / PRIOR APPLICATION NUMBER: 60/090690
      / PRIOR FILING DATE: 1998-06-25
      / PRIOR APPLICATION NUMBER: 60/090694
      / PRIOR FILING DATE: 1998-06-25
      / PRIOR APPLICATION NUMBER: 60/090695
      / PRIOR FILING DATE: 1998-06-25
      / PRIOR APPLICATION NUMBER: 60/090696
      / PRIOR FILING DATE: 1998-06-25
      / PRIOR APPLICATION NUMBER: 60/090862
      / PRIOR FILING DATE: 1998-06-26
      / PRIOR APPLICATION NUMBER: 60/090863
      / PRIOR FILING DATE: 1998-06-26
      / PRIOR APPLICATION NUMBER: 60/091360
      / PRIOR FILING DATE: 1998-07-01
      / PRIOR APPLICATION NUMBER: 60/091478
      / PRIOR FILING DATE: 1998-07-02
      / PRIOR APPLICATION NUMBER: 60/091544
      / PRIOR FILING DATE: 1998-07-01
      / PRIOR APPLICATION NUMBER: 60/091519
      / PRIOR FILING DATE: 1998-07-02
      / PRIOR APPLICATION NUMBER: 60/091626
      / PRIOR FILING DATE: 1998-07-02
      / PRIOR APPLICATION NUMBER: 60/091633
      / PRIOR FILING DATE: 1998-07-02
      / PRIOR APPLICATION NUMBER: 60/091978
      / PRIOR FILING DATE: 1998-07-07
      / PRIOR APPLICATION NUMBER: 60/091982
      / PRIOR FILING DATE: 1998-07-07
      / PRIOR APPLICATION NUMBER: 60/092182
      / PRIOR FILING DATE: 1998-07-09
    /
Alignment Scores:
Pred. No.:          1.38           Length:        2300
Score:              84.00           Matches:         49
Percent Similarity: 38.25%          Conservative:   34
Best Local Similarity: 22.58%       Mismatches:     71
Query Match:        8.42%           Indels:         63
DB:                  9             Gaps:            10
US-09-813-775C-2 (1-193) x US-09-969-734-137 (1-2200)
QY      3 ValHisGLucysProAlatIPLeuTriPLeuLeuSenSerLeuDeuSenRleuProLeugly 22
      : ::::: ||| ::::::::::||| ::::::||| ::::: |||||||||
Db      898 CTCATCATCATCATCTCCCTCTCTTGGGGCTTCCTCGTGCTGTACCTCACAGGT 957
QY      23 Leupro---ValLeuGIyAlAProPIroARLeuLIEcysAsperArgValLeugluarq 41
      :| | ::::: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      958 CTCGCCCGCATGTTCTCCGCCTACTGGGAACTCTGTATCAAGCCCGGCTGTGGAAAC 1017
QY      42 TyrLeuLeuGlualALySGlualGLuAsnIIEthThrcLYcSaLaGIuhIsCysSer 61
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1018 -----CTGGAG-----GAACCACCTGTACTCTCACCCCTTTAGAGAGCAGCC 1055
QY      62 LeuAnsgLuAsnIIEthrValProAsPTPhLyVaAsnPhetYrAlaTrPLysArGMet 81
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1060 CTGACCCCGCAGAGATTGTAATCTACTACTTC-----TGCTAGGCTGGCCTTTA 1104

```

```
OY 82 GluVal-----83
      :::::
Db 1105 GACATGAGCTGCTACACAGACAGTCTGCTCTGCAGACACAGAGGCTCTGTGAG 1164
OY 84 ArgGlnGlnAlaValGluValIrrGln-----GlyLeuAlaLeuLeuSer 98
      :::::
Db 1165 AAGAGCGGAGGCTTCACGCTGGCAACGAACTGGGCTACCCCTGCTATGCTGTGC 1224
OY 99 GluAlaValLeuArgGlyGlnAlaLeuValAsnSer-SerGlnProTrpGluProLe 118
      |||||
Db 1225 TTGCTGTGCTGCTGAGGCGCTGTGCTGCTATTCGCGCATCCACATCCCTGC-----1276
OY 118 uGlnLeuHisValAspLysAlaValSerGlyLeuArgSerLeuThrLeuLeuArgAl 138
      :::::
Db 1277 -----AGCTGCTCATCGATGAGG-----1294
OY 138 aleuGlyAlaGlnLysGluAlaLeuSerProProAspAlaAlaSerAlaAlaProLeuAr 158
      |||||
Db 1295 -CTGCCATGCCCGGAGGCAATGCCAGGCTACCTCTAGGCCAGGCTTCCTTCGAAGCTG 1353
OY 158 gThrIleThrAlaAspThrPheArgLysLeuPheArgValTyrSerAsnPhe-----175
      :::::
Db 1354 GGCTCCTTGGTCCCTCA-----TTCAGTTGTACTCATCTTTTACCTAATG 1401
OY 176 -----LeuArgGlyLysLeuLysLeuTyrThrGlyGluAlaCys 188
      |||||
Db 1402 GTGTCTCAGTTGTGGGCTTATAGCTCTCCACTCTCCGAGGCTGC 1450
```

Search completed: January 8, 2003, 07:52:35
Job time : 70.4596 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 8, 2003, 05:14:09 ; Search time 43.546 Seconds
(without alignments)
1359.222 Million cell updates/sec

Title: US-09-813-775c-2

Perfect score: 998
Sequence: 1 MGVHCPAMWLLSLSLP.....NFLRCKLKLTYGCACTGDR 193

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame.p2n.model -DEV=xlh
-O=/cgn2.1/uspro.spool/us0981377/runat.07012003.153121.23808/app_query.fasta.1.718
-DB=Issued_Patents_NA -QPMF=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human4.0.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pte -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09813775.@cgn.1.18.@runat.07012003.153121.23808 -NCPG=3 -ICPU=3
-NO_XIPXY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMER=120
-MARN_TIMER=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: Issued_Patents_NA.*
- 2: /cgn2.6/ptodata/1/ina/5A.COMB.seq.*
- 3: /cgn2.6/ptodata/1/ina/5B.COMB.seq.*
- 4: /cgn2.6/ptodata/1/ina/6A.COMB.seq.*
- 5: /cgn2.6/ptodata/1/ina/6B.COMB.seq.*
- 6: /cgn2.6/ptodata/1/ina/Backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	991	99.3	725	5 PCT-US94-04361-21 Sequence 21, Appl
2	991	99.3	788	2 US-08-883-795A-35 Sequence 35, Appl
3	991	99.3	823	2 US-08-785-750-1 Sequence 1, Appl
4	991	99.3	823	4 US-09-205-337-1 Sequence 1, Appl
5	991	99.3	1011	2 US-08-750-128-12 Sequence 12, Appl
6	899.5	90.1	5107	4 US-08-910-647-3 Sequence 3, Appl
7	899.5	90.1	5107	4 US-09-620-925-3 Sequence 3, Appl
8	871.5	87.3	681	5 PCT-US94-04361-22 Sequence 22, Appl
9	846	84.8	1239	4 US-09-366-009-39 Sequence 39, Appl
10	843	84.5	508	1 US-08-318-193-69 Sequence 69, Appl
11	821	82.3	945	2 US-08-750-128-7 Sequence 7, Appl
12	821	82.3	969	2 US-08-750-128-4 Sequence 4, Appl

13	792.5	79.4	681	5 PCT-US94-04361-27 Sequence 27, Appl
14	769.5	77.1	688	5 PCT-US94-04361-26 Sequence 26, Appl
15	767	76.9	678	5 PCT-US94-04361-24 Sequence 24, Appl
16	759.5	76.1	687	5 PCT-US94-04361-25 Sequence 25, Appl
17	755	75.7	679	5 PCT-US94-04361-23 Sequence 23, Appl
18	732.5	73.4	3601	3 US-09-017-631-23 Sequence 23, Appl
19	732.5	73.4	3602	2 US-08-883-795A-33 Sequence 33, Appl
20	732.5	73.4	3602	2 US-09-018-138-1 Sequence 1, Appl
21	178	17.8	286	5 PCT-US94-04361-10 Sequence 10, Appl
22	148	14.8	1302	3 US-08-932-823A-1 Sequence 1, Appl
23	134.5	13.5	268	5 PCT-US94-04361-17 Sequence 17, Appl
24	134.5	13.5	269	5 PCT-US94-04361-16 Sequence 16, Appl
25	133.5	13.4	290	5 PCT-US94-04361-11 Sequence 11, Appl
26	130.5	13.1	274	5 PCT-US94-04361-15 Sequence 15, Appl
27	123	12.3	288	5 PCT-US94-04361-13 Sequence 13, Appl
28	120.5	12.1	280	5 PCT-US94-04361-14 Sequence 14, Appl
29	114	11.4	283	5 PCT-US94-04361-19 Sequence 19, Appl
30	112	11.2	283	5 PCT-US94-04361-18 Sequence 18, Appl
31	106	10.6	65	2 US-08-883-795A-24 Sequence 24, Appl
32	105	10.5	58	2 US-08-883-795A-6 Sequence 6, Appl
33	105	10.5	58	2 US-08-883-795A-22 Sequence 22, Appl
34	101	10.1	59	2 US-08-883-795A-4 Sequence 4, Appl
35	100.5	10.1	286	5 PCT-US94-04361-12 Sequence 12, Appl
36	97	9.7	56	2 US-08-883-795A-5 Sequence 5, Appl
37	97	9.7	61	2 US-08-883-795A-8 Sequence 8, Appl
38	96.5	9.7	605	1 US-08-388-779A-3 Sequence 3, Appl
39	96.5	9.7	605	1 US-08-591-070A-3 Sequence 3, Appl
40	96.5	9.7	605	2 US-08-927-855-3 Sequence 3, Appl
41	96.5	9.7	1059	3 US-08-471-045-57 Sequence 57, Appl
42	96.5	9.7	1059	3 US-08-469-712A-57 Sequence 57, Appl
43	96.5	9.7	1059	4 US-08-446-871-57 Sequence 57, Appl
44	96.5	9.7	1059	4 US-08-468-910-57 Sequence 57, Appl
45	96.5	9.7	1059	4 US-08-761-907-57 Sequence 57, Appl

ALIGNMENTS

RESULT 1
Sequence 21, Application PC/TUS9404361
GENERAL INFORMATION:
APPLICANT: Brigham and Women's Hospital
APPLICANT: 75 Francis Street
APPLICANT: Boston, MA 02115
APPLICANT: Bunn, H. Franklin
APPLICANT: Wen, Danyl
TITLE OF INVENTION: Erythropoietin Mutains With Enhanced
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04361
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/049,802
FILING DATE: 21-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Cimbalia, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0627.336PC01

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 725 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: both
 PCT-US94-04361-21

Alignment Scores:

Pred. No.:	2,596-119	Length:	725
Score:	991.00	Matches:	192
Percent Similarity:	99.48%	Conservative:	0
Best Local Similarity:	99.48%	Mismatches:	1
Query Match:	99.30%	Indels:	0
DB:	5	Gaps:	0

US-09-813-775c-2 (1-193) x PCT-US94-04361-21 (1-725)

QY 1 MetGlyValHisGluCysProAlaTrpLeuTrpLeuLeuSerLeuLeuSerLeuPro 20
 DB 11 ATGGGGGTCACGAATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 70
 QY 21 LeuGlyLeuProValLeuGlyAlaProProAlaGlyLeuIleCysAspSerArgValLeuGlu 40
 DB 71 CTGGGGCTCCAGTCTGCTGGGCGCCGCCACACGCTCATCTGTGACAGCGAGCTGTGGAG 130
 QY 41 ArgTyrLeuLeuGluAlaLeuGluAlaGluAlaGluAlaGluAlaGluAlaGluAla 60
 DB 131 AGGTACCTTGTGAGGCGCCAGAGAGCGCCGGAATATCAGACGCGCTGTCTGACACTGC 190
 QY 61 SerLeuAsnGluAsnIleThrValProAspThrLysValAsnPheTyrAlaTrpLysArg 80
 DB 191 ACCTGAATGAGAAATATCACTGCTCCAGACACCAAGTTAATTTCATGCTCGGAGAGAG 250
 QY 81 MetGluValArgGlnAlaValAlaGluValITrpgInGlyLeuAlaLeuLeuSerGluAla 100
 DB 251 ATGGAGGCTGGGCGACGAGCGCTAGAGAGCTGGGAGGCGCTGCTGCTGCGGAGAGCT 310
 QY 101 ValLeuArgGlyGlnAlaLeuLeuValAsnSerSerGlnProTrpGlnProLeuGlnLeu 120
 DB 311 CTCCTCGGGGCGGAGGCGCTGTTGGTCAACTCTTCCAGCCGCTGGGAGCGCTGCGACTG 370
 QY 121 HisValAspLysAlaValSerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGly 140
 DB 371 CATGTGATTAAGCCGTCAGTGGCTTCGACGCTCACCACACCTGCTGCGGCTGCGGAG 430
 QY 141 AlaGlnLysGluAlaIleSerProProAspAlaIleSerAlaAlaProLeuArgThrIle 160
 DB 431 GCCCAAGAAAGAGCCATCTCCCTCCAGATGGGCTCAGCTGCTCACTCCGAAACAAATC 490
 QY 161 ThrAlaAspThrPheArgLysLeuPheArgValTyrSerAsnPheLeuArgGlyLysLeu 180
 DB 491 ACTGCTGACACTTCCGCAACTCTTCCAGTCTACTCCAAATTTCTCTCCGGGAAAGCTG 550
 QY 181 LysLeuTyrThrGlyGluAlaCysArgThrGlyAspArg 193
 DB 551 AAGCTGTACACAGGGAGGCTGTGACGAGACAGGGAGACA 589

RESULT 2

US-08-883-795A-35
 Sequence 35, Application US/08883795A

GENERAL INFORMATION:
 APPLICANT: Delcive, Genevieve
 APPLICANT: Awang, Gregor
 TITLE OF INVENTION: Recombinant DNA Molecules and Expression
 TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
 NUMBER OF SEQUENCES: 39
 CORRESPONDENCE ADDRESS:
 ADDRESS: BERESKIN & PARR

STREET: 40 King Street West
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada
 ZIP: M5H 3Y2

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/883,795A

FILING DATE: 27-JUN-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Gravelle, Micheline

REGISTRATION NUMBER: 40,261

REFERENCE/DOCKET NUMBER: 7841-062

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 364-7311

TELEFAX: (416) 361-1398

INFORMATION FOR SEQ ID NO: 35:

SEQUENCE CHARACTERISTICS:

LENGTH: 788 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

IMMEDIATE SOURCE:

CLONE: EPOLong

US-08-883-795A-35

Alignment Scores:

Pred. No.:	2,966-119	Length:	788
Score:	991.00	Matches:	192
Percent Similarity:	99.48%	Conservative:	0
Best Local Similarity:	99.48%	Mismatches:	1
Query Match:	99.30%	Indels:	0
DB:	2	Gaps:	0

US-09-813-775c-2 (1-193) x US-08-883-795A-35 (1-788)

QY 1 MetGlyValHisGluCysProAlaTrpLeuTrpLeuLeuSerLeuLeuSerLeuPro 20
 DB 80 ATGGGGGTCACGAATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 139
 QY 21 LeuGlyLeuProValLeuGlyAlaProProAlaGlyLeuIleCysAspSerArgValLeuGlu 40
 DB 140 CTGGGGCTCCAGTCTGCTGGGCGCCGCCACACGCTCATCTGTGACAGCGAGCTGTGGAG 199
 QY 41 ArgTyrLeuLeuGluAlaLeuGluAlaGluAlaGluAlaGluAlaGluAlaGluAla 60
 DB 200 AGGTACCTTGTGAGGCGCCAGAGAGCGCCGGAATATCAGACGCGCTGTCTGACACTGC 259
 QY 61 SerLeuAsnGluAsnIleThrValProAspThrLysValAsnPheTyrAlaTrpLysArg 80
 DB 260 ACCTGAATGAGAAATATCACTGCTCCAGACACCAAGTTAATTTCATGCTCGGAGAGAG 319
 QY 81 MetGluValArgGlnAlaValAlaGluValITrpgInGlyLeuAlaLeuLeuSerGluAla 100
 DB 320 ATGGAGGCTGGGCGACGAGCGCTAGAGAGCTGGGAGGCGCTGCTGCTGCGGAGAGCT 379
 QY 101 ValLeuArgGlyGlnAlaLeuLeuValAsnSerSerGlnProTrpGlnProLeuGlnLeu 120
 DB 380 CTCCTCGGGGCGGAGGCGCTGTTGGTCAACTCTTCCAGCCGCTGGGAGCGCTGCGACTG 439
 QY 121 HisValAspLysAlaValSerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGly 140
 DB 440 CATGTGATTAAGCCGTCAGTGGCTTCGACGCTCACCACACTCTGCTGCGGCTGCGGAG 499
 QY 141 AlaGlnLysGluAlaIleSerProProAspAlaIleSerAlaAlaProLeuArgThrIle 160

|||||
Db 500 GCCCAGAGGAGCATTCTCCCTCCAGATGCGCTCAGCTGCTCCACTCCGAAACATC 559
Qy 161 ThrAlaSpThrPheArgLysLeuPheArgValTyrSerAsnPheLeuArgGlyLysLeu 180
Db 560 ACAGCTGACACTTCCGCAAACTCTCCGAGTCTACTCCATTCTCCGCGGAAAGCTG 619
Qy 181 LysLeuTyrThrGlyLysAlaCysArgThrGlyAspArg 193
Db 620 AACCTGTACACAGGAGGAGCTCCAGAGACAGCGGACAGA 658
RESULT 3
US-08-785-750-1
Sequence 1, Application US/08785750
Patent No. 5846528
GENERAL INFORMATION:
APPLICANT: PODSAKOFF, GREGORY M.
APPLICANT: KURTZMAN, GARY J.
TITLE OF INVENTION: METHODS OF TREATING ANEMIA USING
TITLE OF INVENTION: RECOMBINANT ADENO-ASSOCIATED VIRUS VARIANTS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROBINS & ASSOCIATES
STREET: 90 MIDDLEFIELD ROAD, SUITE 200
CITY: MEMLO PARK
STATE: CA
COUNTRY: USA
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/785,750
FILING DATE: 16-JAN-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/588,355
FILING DATE: 18-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 0800-0009.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 325-7812
TELEFAX: (415)325-7823
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 823 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-785-750-1
Alignment Scores:
Pred. No.: 3,17e-119 Length: 823
Score: 991.00 Matches: 192
Percent Similarity: 99.488 Conservative: 0
Best Local Similarity: 99.488 Mismatches: 1
Query Match: 99.308 Indels: 0
Gaps: 0
US-09-813-775C-2 (1-193) x US-08-785-750-1 (1-823)
Qy 1 MetGlyValHisIleuGlyProAlaIleuPheLeuLeuLeuSerLeuLeuSerLeuPro 20
Db 52 ATGGGGGTGACCAAAATGCTCGCTGCTGCTCTCTCTGCTCCCTGCTGCTCCCT 111
Qy 21 LeuGlyLeuProValLeuGlyAlaProProArgLeuIleCysAspSerArgValLeuGlu 40
Db 112 CTGGGCTCCAGTCTCTGGGCGCCCAACAGCGCTCATCTGTGACACAGCGAGTCTGGAG 171

Qy 41 ArgTyrLeuLeuGluAlaLysGluAlaGluAsnIleThrThrGlyCysAlaGluHisCys 60
Db 172 AGGTACCTCTTGGAGGCGCAAGAGCGCCGACAGATATACACAGCGGCTGTGCTGAACACTGC 211
Qy 61 SerLeuAsnGluAsnIleThrValProAspThrLysValAsnPheTyrAlaIlePlyAspArg 80
Db 232 AGCTTGAATGAGATACATGCTCCAGACCAAAAGTTATTTCTATGCTGGAAGAGG 291
Qy 81 MetGlyValArgGlnGlnAlaValAlaGluValTyrGlnGlyLeuAlaLeuSerGluAla 100
Db 292 ATGAGAGTCCGGGACGACAGGCGCTGAGAGTCTGGCAGGCGCTGCGCTGTGGAAAGCT 351
Qy 101 ValLeuArgGlyGlnAlaLeuLeuValAsnSerSerGlnProTyrGluProLeuGlnLeu 120
Db 352 GTCTCGGGGCGCAGGCGCTGTGTGTAACCTCTCCAGCGGAGGAGCGCCCTCGACAGCTG 411
Qy 121 HisValAspLysAlaValSerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGly 140
Db 412 CATGTGATTAAGCCGTCATGCTCCGACGCTCACCACTCTGCTTCGGGCTCTGGGA 471
Qy 141 AlaGlnLysGluAlaIleSerProProAspAlaAlaSerAlaIleProLeuArgTyrIle 160
Db 472 GCCCAGAGGAGCATTCTCCCTCCAGATGCGGCTCAGCTGCTCCACTCCGAAACATC 531
Qy 161 ThrAlaSpThrPheArgLysLeuPheArgValTyrSerAsnPheLeuArgGlyLysLeu 180
Db 532 ACTGCTGACACTTCCGCAAACTCTCCGAGTCTACTCCATTCTCCGCGGAAAGCTG 591
Qy 181 LysLeuTyrThrGlyLysAlaCysArgThrGlyAspArg 193
Db 592 AACGTGTACACAGGAGGAGCTCCAGAGACAGCGGACAGA 630
RESULT 4
US-09-205-337-1
Sequence 1, Application US/09205337
Patent No. 6325998
GENERAL INFORMATION:
APPLICANT: PODSAKOFF, GREGORY M.
APPLICANT: KURTZMAN, GARY J.
TITLE OF INVENTION: METHODS OF TREATING ANEMIA USING
TITLE OF INVENTION: RECOMBINANT ADENO-ASSOCIATED VIRUS VARIANTS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROBINS & ASSOCIATES
STREET: 90 MIDDLEFIELD ROAD, SUITE 200
CITY: MEMLO PARK
STATE: CA
COUNTRY: USA
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/205,337
FILING DATE: 04-Dec-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/785,750
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: MCCracken, THOMAS P.
REGISTRATION NUMBER: 38,548
REFERENCE/DOCKET NUMBER: 0800-0009.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 325-7812
TELEFAX: (415)325-7823
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 823 base pairs
TYPE: nucleic acid

US-09-366-009-39
Sequence 39, Application US/09366009
Patent No. 6426042
GENERAL INFORMATION:
APPLICANT: Asada, Kiyozo
Uemori, Takashi
Koyama, No. 6426042uto
Hashino, Kimikazu
Kato, Ikunoshin
TITLE OF INVENTION: METHOD FOR GENE TRANSFER INTO TARGET
CELLS WITH RETROVIRUS
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: WEISER & ASSOCIATES
STREET: 230 South Fifteenth Street, Suite 500
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/366,009
FILING DATE: 02-Aug-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/809,156
FILING DATE: <Unknown>
APPLICATION NUMBER: JP 294382/1995
FILING DATE: 13-NOV-1995
APPLICATION NUMBER: JP 051847/1996
FILING DATE: 08-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Weiser, Gerard J.
REGISTRATION NUMBER: 19,763
REFERENCE/DOCKET NUMBER: 977,6507P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-875-8383
TELEFAX: 215-875-8394
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 1239 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-09-366-009-39
Alignment Scores:
Pcted. No.: 4,22e-100 Length: 1239
Score: 846.00 Matches: 165
Percent Similarity: 99.40% Conservative: 1
Best Local Similarity: 98.80% Mismatches: 1
Query Match: 84.77% Indels: 0
DB: 4 Gaps: 0
US-09-813-775C-2 (1-193) x US-09-366-009-39 (1-1239)
QY 27 G1YAlAProProlArgLeuIleCysAspSerArgValLeuGluArgTyrLeuLeuGluAla 46
|||||
Db 694 GGTGCCCGACACGCCCTCATCTGTGACAGCGAGCTCTGCAGAGGTACCTCTTGAGAGGCC 753
QY 47 LysGluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIle 66
|||||
Db 754 AAGGAGGCCGAGATATACGACGAGGCTGTGTGACACACTGCACTGGAATGACATATAC 813
QY 67 ThValProAspThrIysValAsnPhetYrAlaTrpIysArgMetGluValArgGlnGln 86
|||||

Db 814 ACTGTCCACAGACCAAGTTAATTCTATGCTCGAAGAGAGTGAAGCTGGCGACAGC 873
QY 87 AlAValGluValTrpGlnGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAla 106
|||||
Db 874 GCCGTAGAGTCTGGCAGGGGCTGGCCCTGCTGTGCGAAGCTGTGCGGGGGCAGAGCC 933
QY 107 LeuLeuValAsnSerSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaVal 126
|||||
Db 934 CTGTGTGTCAACTCTTCCAGCCGCTGGAGCCCTGCAGCTGATGTGATTAAGCCGTC 993
QY 127 SerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGluAlaIle 146
|||||
Db 994 AGTGGCTTCGACACCTCCACCACTCTCTGGGCTCTGGAGCCGAGGAAGGAGCCATC 1053
QY 147 SerProAspAlaAlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArg 166
|||||
Db 1054 TCCCTTCAGATGGGGCTCAGCTGCTCCACCTCGAACAATCACTGCTGACACTTTCGC 1113
QY 167 LysLeuPheArgValTyrSerAsnPhelLeuArgGlyLysLeuLysLeuTyrThrGlyGlu 186
|||||
Db 1114 AAACCTTCGAGACTACTCCAAATTTCCTCCGGGAAGCTGAAGCTGTACACAGGGAG 1173
QY 187 AlaCysArgThrGlyAspArg 193
Db 1174 GCTGTGAGACAGGAGGAGAGA 1194
RESULT 10
US-08-318-193-69
Sequence 69, Application US/08318193
Patent No. 5641663
GENERAL INFORMATION:
APPLICANT: GARVIN, Robert T.
TITLE OF INVENTION: AN EXPRESSION SYSTEM FOR THE SECRETION
TITLE OF INVENTION: OF BIOACTIVE HUMAN GRANULOCYTE MACROPHAGE COLONY
STIMULATING FACTOR (GM-CSF) AND OTHER HETEROLOGOUS
NUMBER OF SEQUENCES: 91
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: Virginia
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/318,193
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,314
FILING DATE:
APPLICATION NUMBER: US 07/224,568
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 18740/116 CACO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 508 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: linear
MOLECULE TYPE: other nucleic acid;

DESCRIPTION: Synthetic DNA oligonucleotide
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 5..502
 US-08-318-193-69

Alignment Scores:
 Pred. No.: 2,46e-100 Length: 508
 Score: 843.00 Matches: 165
 Percent Similarity: 99.40% Conservative: 0
 Best Local Similarity: 99.40% Mismatches: 1
 Query Match: 84.47% Indels: 0
 DB: 1 Gaps: 0

US-09-813-775c-2 (1-193) x US-08-318-193-69 (1-508)

QY 28 AlaProProArgLeuIleCysAspSerArgValLeuGluArgTyrLeuGluAlaLys 47
 DB 5 GCGCCCCCGGCTCATCTGCGACACCGCGCTCTCGAGCGGTAAGCTGCTCGAGCGCAG 64
 QY 48 GluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIleThr 67
 DB 65 GAGGCGGAGAAATATCATCGAGGCGGCTGCGGAGCACTGCTCCCTCAACGAGAAATCAGC 124
 QY 68 ValProAspThrLysValAsnPhenylAlaTyrPlyArgMetGluValArgGlnGlnAla 87
 DB 125 GTCCCGACACAGGTCACTTCTACAGCTGAGACGCGATGAGGTGGGCGCAGCGGCG 184
 QY 88 ValGluValTrpGlnGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlnAlaLeu 107
 DB 185 GTGCAAGTGTGGCAGGCGCTCGCGCTCTCTCGAGGCGGCTCTCGGCGCGCGGCTCTC 244
 QY 108 LeuValAsnSerSerGlnProTrrpGluProLeuGlnLeuHisValAspLysAlaValSer 127
 DB 245 CTGGTGAACCTGCTCCAGCGCGTGGAGCGCGCTCCAGCTCAGCTGACAGCAAGGCGCTCTC 304
 QY 128 GlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGluAlaGlnLysGluAlaLeuSer 147
 DB 305 GGGCTGGTCTCTGACCGAGCTGTGCTGCGCTCTGCGGCGCGGCGCGGCGGCGGCTCTC 364
 QY 148 ProProAspAlaAlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLys 167
 DB 365 CCGCGGAGCGCGCGCGCGCGCGCGCGCTGCGGAGCATATCAGCGGCGGCGGCGGCGGCGG 424
 QY 168 LeuPheArgValTyrSerAsnPhenylArgGlyLysLeuLysLeuTyrThrGlyGluAla 187
 DB 425 CTGTCGCGGCTACTCGAACTTCTGCGGCGGAGAACTGAACTGATACACCGCGGCGGCGG 484
 QY 188 CysArgThrGlyAspArg 193
 DB 485 TGCCGACGCGGCGGCGGCGG 502

RESULT 11

US-08-750-128-7

Sequence 7, Application US/08750128

Patent No. 5916773

GENERAL INFORMATION:

APPLICANT: MELE Antonio,

APPLICANT: DE SANTIS Rita,

APPLICANT: CARLONI Cristina,

APPLICANT: COSCARIELLA Annamaria

TITLE OF INVENTION: Hybrid molecule of formula GM-CSF-L-EPO or

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: LOWE, PRICE, LEBLANC & BECKER

STREET: 99 Canal Center Plaza, Suite 300

CITY: Alexandria

STATE: VIRGINIA

COUNTRY: USA

ZIP: 22314

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk, 3.50 inch.

COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentia Release #1.0, Version #1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/750,128
 FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: IT FI 94 A 000106

FILING DATE: 27-MAY-1994

ATTORNEY/AGENT INFORMATION:

NAME: Robert L. Price

REGISTRATION NUMBER: 22,685

REFERENCE/DOCKET NUMBER: 2879-007

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-684-1111

TELEFAX: 703-684-1124

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 945 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

HYPOTHETICAL: NO

US-08-750-128-7

Alignment Scores:
 Pred. No.: 4.83e-97 Length: 945
 Score: 821.00 Matches: 162
 Percent Similarity: 97.01% Conservative: 0
 Best Local Similarity: 97.01% Mismatches: 5
 Query Match: 82.26% Indels: 0
 DB: 2 Gaps: 0

US-09-813-775c-2 (1-193) x US-08-750-128-7 (1-945)

QY 27 GlyAlaProProArgLeuIleCysAspSerArgValLeuGluArgTyrLeuGluAla 46
 DB 442 GGGGCGCGGCGCTCACTCATCTGTGACAGCGAGTCTCGAGAGGTAAGCTCTGGAGGCG 501
 QY 47 LysGluAlaGluAsnIleThrThrGlyCysAlaGluHisCysSerLeuAsnGluAsnIle 66
 DB 502 AAGGAGCGCGAATATCATCGAGGCGGCTGCTGAACACTGCACTGCAATGAGAAATATC 561
 QY 67 ThrValProAspThrLysValAsnPhenylArgTyrPlyArgMetGluValArgGln 86
 DB 562 ACTGTCGCCAGACACCAAGTTAATTTCTATGCTTGGAGAGGATGAGGTGCGGCGAG 621
 QY 87 AlaValGluValTrpGlnGlyLeuAlaLeuLeuSerGluAlaValLeuArgGlyGlnAla 106
 DB 622 GCCGTAGAAGTGTGCGAGGCGCTGCGCTGCTGCGAAGCTGTCTGCGGCGGCGGCGGCG 681
 QY 107 LeuLeuValAsnSerSerGlnProTrrpGluProLeuGlnLeuHisValAspLysAlaVal 126
 DB 682 CTGTTGGTCAACTCTCCAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 741
 QY 127 SerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGluAlaGlnLysGluAlaLeu 146
 DB 742 AATGCGCTTCCGAGCTTCACTGCTGCTGCGGCTTGGGAGCCGAGAGGAGGAGGAGGAG 801
 QY 147 SerProProAspAlaAlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArg 166
 DB 802 TCCTCCGAGATGCGGCGCTGAGCTGCTGCTGCGGCTTGGGAGCCGAGAGGAGGAGGAG 861
 QY 167 LysLeuPheArgValTyrSerAsnPhenylArgGlyLysLeuLysLeuTyrThrGlyGlu 186
 DB 862 AAACCTTCGAGTCTACTCCCAATTTCTCCGCGGAGAACTGAACTGAACTGATACACAGGAG 921
 QY 187 AlaCysArgThrGlyAspArg 193
 DB 922 GCTCGAGACAGGCGGAGCAGA 942

RESULT 12
US-08-750-128-4
Sequence 4, Application US/08750128
Patent No. 5916773
GENERAL INFORMATION:
APPLICANT: MELE Antonio,
DE SANTIS Rita,
APPLICANT: CARLOTTI Cristina,
APPLICANT: COSCARIELLA Annamaria
TITLE OF INVENTION: Hybrid molecule of formula GM-CSF-L-EPO or
TITLE OF INVENTION: EPO-L-GM-CSF for hematopoietic stimulation
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: LOWE, PRICE, LeBLANC & BECKER
STREET: 99 Canal Center Plaza, Suite 300
CITY: Alexandria
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22314
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk, 3.50 inch.
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/750,128
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IT FI 94 A 000106
FILING DATE: 27-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Robert L. Price
REGISTRATION NUMBER: 22,685
REFERENCE/DOCKET NUMBER: 2879-007
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-684-1111
TELEFAX: 703-684-1124
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 969 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
US-08-750-128-4
Alignment Scores:
Pred. No.: 5.03e-97 Length: 969
Score: 821.00 Matches: 162
Percent Similarity: 97.01% Conservative: 0
Best Local Similarity: 97.01% Mismatches: 5
Query Match: 82.26% Indels: 0
Gaps: 0
US-09-813-775C-2 (1-193) x US-08-750-128-4 (1-969)
QY 27 GLYALAPROPRIOARGLEULECYSASPSEARGVALLLEAGUARGTYRLEULEUVALA 46
Db 466 GGGGGGGGGGGCTCACTATCTGTGACAGCCGAGTCTGGAGAGGTACTTGGAGGCC 525
QY 47 LYSGLUALAGLUASNLEIETHTHRLCYSLAAGLUHISCYSSERLEUASNGLUASNLE 66
Db 526 AAGGAGCGCCGGAATATACGACGCGGCTGTGTGAACAACCTGAGTTCAGATGACATATC 585
QY 67 THRVAlPROASPThRLYSVALASNPhETyRALATrPLYSAlRGMeTGLUVALArgLInGLn 86
Db 586 ACTGTCCAGACACCAAGATTAAATTCTATGCTCGAAGAGAGATGAGGTCCGCGCAGCAG 645
QY 87 ALAVALGLUVALTRPGLInGLYLeUAlALEULeUSERGLUAlAVALLeUArgLGLYInAlA 106
Db 646 GCCGTAGAAGTCTGGCAGGGCTGGCCCTGCTGTGGAAGCTGTCTGGGGGGCCAGGCC 705

QY 107 LEULeUVALASNSErSerGLNPrOTrPGluProlEUGInLEUHLsVALAsPLYSAlAVAL 126
Db 706 CTCTTGCTCAACTCTTCCACGCGGAGGCCCTCGACAGCTGCATGATTAACCCCTC 765
QY 127 SERGLYLeUArgSERLeUTHrThRLLeUeUArgAlALEUGLYAlAGLInGLYSLAlle 146
Db 766 AATGGCTCTTCGACGCTCACCTCTGCTTCGGGCTGTGGAGCCAGAGAACGACCATC 825
QY 147 SERPRoPAsPAlAlAsERAlAlAProlEUAArgThRLIEThRALAsPThrPhEARg 166
Db 826 TCCCTCCAGATGGGCGCTCAGCTGCTCCACTCCGACACATACGTGACACTTCCGC 885
QY 167 LYSLeUPhEARgVALTYrSErASNPhELeUArgGLYLSLeUYSLeUyTrThRLGLYLU 186
Db 886 AAACCTCTCCGAGCTACTGCCAATTTCCGCCGGGGAAGAGCTGAAGCTGACACAGGGAG 945
QY 187 ALAcYsArgThGLYASPARg 193
Db 946 GCCTGACGACAGGGGACAGA 966
RESULT 13
PCT-US94-04361-27
Sequence 27, Application PC/TUS9404361
GENERAL INFORMATION:
APPLICANT: Brigham and Women's Hospital
APPLICANT: 75 Francis Street
APPLICANT: Boston, MA 02115
APPLICANT: Bunn, H. Franklin
APPLICANT: Wen, Danyi
APPLICANT: Showers, Mark O.
TITLE OF INVENTION: Erythropoietin Mutains With Enhanced
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04361
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/049,802
FILING DATE: 21-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Cimdala, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0627.336PC01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
FAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 681 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: both
PCT-US94-04361-27
Alignment Scores:
Pred. No.: 1.44e-93 Length: 681
Score: 792.50 Matches: 159
Percent Similarity: 88.36% Conservative: 8
Best Local Similarity: 84.13% Mismatches: 21

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 00:53:52 ; Search time 51.61 Seconds
(without alignments)
498.302 Million cell updates/sec

Title: US-09-813-775C-2
Perfect score: 998
Sequence: 1 MGVECPAMWLLSLSLP.....NLRGKIKLYTGACRTGDR 193

Scoring table: BLOSUM62
Gap 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A_Geneseq_101002.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
24: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	998	100.0	194	22	AA34977
2	992	99.4	193	22	AA34979
3	991	99.3	193	6	AA50300
4	991	99.3	193	8	AA70256
5	991	99.3	193	15	AA65499
6	991	99.3	193	16	AA81982
7	991	99.3	193	16	AA81137
8	991	99.3	193	17	AA84141
9	991	99.3	193	16	AA98397
10	991	99.3	193	21	AA94550

11	991	99.3	193	21	AA93638	Amino acid sequenc
12	991	99.3	193	21	AA93704	Human non-91cosyl
13	991	99.3	193	21	AA94398	Human erythropoiet
14	991	99.3	193	22	AA88573	Human erythropoiet
15	991	99.3	193	22	AA834978	Human erythropoiet
16	991	99.3	193	23	AAE15341	Human erythropoiet
17	991	99.3	330	13	AA823076	Epo-IL-3 short, re
18	991	99.3	376	20	AA823079	Epo-IL-3 Flex, rec
19	991	99.0	376	20	AA823079	Human erythropoiet
20	988	99.0	193	16	AA60598	Open reading frame
21	988	99.0	193	16	AA874560	Erythropoietin. H
22	987	98.9	193	16	AA81986	Erythropoietin (A1
23	987	98.9	193	21	AA94536	Human erythropoiet
24	987	98.9	193	21	AA94536	Human EPO protein
25	985	98.7	193	18	AA81913	Erythropoietin var
26	985	98.7	193	19	AA858400	Human erythropoiet
27	984	98.6	193	12	AA811859	Erythropoietin ana
28	984	98.6	193	16	AA81983	Erythropoietin (L6
29	983	98.5	193	18	AA814142	Erythropoietin var
30	983	98.5	193	18	AA814144	Erythropoietin var
31	983	98.5	193	19	AA851935	Erythropoietin. H
32	982	98.4	193	16	AA81987	Erythropoietin (A1
33	982	98.4	193	18	AA814145	Erythropoietin var
34	982	98.4	193	18	AA811892	Erythropoietin ana
35	981	98.3	193	16	AA81984	Erythropoietin (L6
36	980.5	98.2	194	19	AA862048	Human erythropoiet
37	980.5	98.2	194	21	AA810654	Human erythropoiet
38	980.5	98.2	193	21	AA94531	Human erythropoiet
39	980	98.2	193	21	AA94531	Human erythropoiet
40	980	98.2	193	21	AA94532	Human erythropoiet
41	979.5	98.1	193	23	AA877902	Amino acid sequenc
42	979	98.1	193	21	AA94534	Human erythropoiet
43	978	98.0	193	16	AA81985	Erythropoietin (A6
44	978	98.0	193	22	AA835014	Chimpanzee erythro
45	977	97.9	193	20	AA873561	Human erythropoiet

ALIGNMENTS

RESULT 1
ID AAB34977 standard; Protein: 194 AA.
XX AAB34977;
AC AAB34977;
XX 28-MAR-2001 (first entry)
DE Chimpanzee erythropoietin SEQ ID NO: 2.
XX Chimpanzee; erythropoietin; EPO; hybridisation probe; gene therapy;
KW mapping; therapeutic agent.
XX
XX Pan sp.
OS WO200068376-A1.
PN 16-NOV-2000.
PD 05-MAY-2000; 2000WO-US12370.
PR 07-MAY-1999; 99US-0307307.
PR 28-MAR-2000; 2000US-0307307.
PA (GETH) GENENTECH INC.
XX Desauvage F, Henner DJ;
PI WPI: 2001-007393/01.
XX N-PSDB; AAC66881.
DR Nucleic acids encoding chimpanzee erythropoietin, useful for treatment
XX of e.g. anemia, also derived proteins, antibodies and modulators -
PT

XX Claim 1; Fig 1; 109pp; English.
 PS The present invention provides the coding and protein sequences of
 CC chimpanzee erythropoietin (EPO). These sequences can be used in gene
 CC therapy, to block the activity of EPO, as hybridisation probes, in
 CC genetic and chromosome mapping and as therapeutic agents.
 XX
 SQ Sequence 194 AA:
 Query Match 100.0%; Score 998; DB 22; Length 194;
 Best Local Similarity 100.0%; Pred. No. 4,8e-100;
 Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGVHECPAWMLLSTLSPGLPVGAPPRILCDSRVLERLYLLEAKENITTGCAEHC 60
 DB 1 MGVHECPAWMLLSTLSPGLPVGAPPRILCDSRVLERLYLLEAKENITTGCAEHC 60
 QY 61 SLNENITVPDTKVFYAMKRMVEYRQOAVEWQGLALSEAVLRGQALLVNSSQPMPEPLQ 120
 DB 61 SLNENITVPDTKVFYAMKRMVEYRQOAVEWQGLALSEAVLRGQALLVNSSQPMPEPLQ 120
 QY 121 HDKAVSGLSLTTLLRALGAQKEAISPDDASAAPLRTITADTFKRLFRVYSNFLRGKL 180
 DB 121 HDKAVSGLSLTTLLRALGAQKEAISPDDASAAPLRTITADTFKRLFRVYSNFLRGKL 180
 QY 181 KLYTGEACRTGDR 193
 DB 181 KLYTGEACRTGDR 193
 RESULT 2
 AAB34979
 ID AAB34979 standard; Protein: 193 AA.
 AC AAB34979;
 DT 27-MAR-2001 (first entry)
 DE Chimpanzee erythropoietin SEQ ID NO: 5.
 XX Chimpanzee; erythropoietin; EPO; hybridisation probe; gene therapy;
 KM mapping; therapeutic agent.
 XX
 OS Pan sp.
 PN WO200068376-A1.
 XX 16-NOV-2000.
 PD 05-MAY-2000; 2000MO-US12370.
 PF 07-MAY-1999; 99US-0307307.
 PR 28-MAR-2000; 2000US-0307307.
 XX (GETH) GENENTECH INC.
 PA Desauvage F, Henner DJ;
 PI WPI: 2001-007393/01.
 DR Nucleic acids encoding chimpanzee erythropoietin, useful for treatment
 PT of e.g. anemia, also derived proteins, antibodies and modulators -
 XX
 PS Disclosure: Fig 3; 109pp; English.
 CC The present invention provides the coding and protein sequences of
 CC chimpanzee erythropoietin (EPO). These sequences can be used in gene
 CC therapy, to block the activity of EPO, as hybridisation probes, in
 CC genetic and chromosome mapping and as therapeutic agents.
 XX
 SQ Sequence 193 AA:

Query Match 99.4%; Score 992; DB 22; Length 193;
 Best Local Similarity 99.5%; Pred. No. 2.2e-99;
 Matches 192; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MGVHECPAWMLLSTLSPGLPVGAPPRILCDSRVLERLYLLEAKENITTGCAEHC 60
 DB 1 MGVHECPAWMLLSTLSPGLPVGAPPRILCDSRVLERLYLLEAKENITTGCAEHC 60
 QY 61 SLNENITVPDTKVFYAMKRMVEYRQOAVEWQGLALSEAVLRGQALLVNSSQPMPEPLQ 120
 DB 61 SLNENITVPDTKVFYAMKRMVEYRQOAVEWQGLALSEAVLRGQALLVNSSQPMPEPLQ 120
 QY 121 HDKAVSGLSLTTLLRALGAQKEAISPDDASAAPLRTITADTFKRLFRVYSNFLRGKL 180
 DB 121 HDKAVSGLSLTTLLRALGAQKEAISPDDASAAPLRTITADTFKRLFRVYSNFLRGKL 180
 QY 181 KLYTGEACRTGDR 193
 DB 181 KLYTGEACRTGDR 193
 RESULT 3
 AAP50300
 ID AAP50300 standard; protein: 193 AA.
 AC AAP50300;
 DT 01-JAN-1980 (first entry)
 DE Human erythropoietin encoded by positive clone (phage lambda-hel)
 DE isolated from human fetal liver gene bank.
 XX
 KM Erythropoietin; red blood cell; erythrocyte; anaemia; blood;
 XX disorder; ss; phage lambda-hel; gene bank.
 OS Homo sapiens.
 PN WO8502610-A.
 XX 20-JUN-1985.
 PD 11-DEC-1984; 84MO-US02021.
 PF 30-NOV-1984; 84US-0675298.
 PR 13-DEC-1983; 83US-0561024.
 PR 21-FEB-1984; 84US-0582185.
 PR 28-SEP-1984; 84US-0655841.
 XX (KIRI-) KIRIN-AMGEN INC.
 PA WPI: 1985-159229/26.
 DR N-PSDB; AAN50347.
 DR New polypeptide having properties of erythropoietin - is prepd.
 PT by cultivation of transformed eucaryotic or procaryotic host
 XX
 PS Disclosure: Page 43; 113pp; English.
 CC Human erythropoietin encoded by a sequence encoded by this phage
 CC lambda-hel is essential for red blood cell formation and is used
 CC for the diagnosis and treatment of blood disorders such as anaemia.
 CC Large amounts of EPO may be obtained using recombinant DNA
 CC techniques in contrast to small amounts obtained from plasma
 CC and urine. This sequence is expressed in E. coli. See also
 CC AAN50345-6, AAN50348-50 and AAP50298-99, AAP50301.
 XX
 SQ Sequence 193 AA:
 Query Match 99.3%; Score 991; DB 6; Length 193;
 Best Local Similarity 99.5%; Pred. No. 2.8e-99;
 Matches 192; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MGVHECPAWMLLSTLSPGLPVGAPPRILCDSRVLERLYLLEAKENITTGCAEHC 60

Db 1 MGVECPAMWMLLSTLSPGLPVIGAPPRLLICDSRVLEERYLLLEAKENITTCGAHC 60
 QY SLNENITVPDTKVFYAMKREMEVROQAVEWOGIALISEAVLRQALLVNSOPWEPLQL 120
 Db 61 SLNENITVPDTKVFYAMKREMEVROQAVEWOGIALISEAVLRQALLVNSOPWEPLQL 120
 QY 121 HDKAVSGLSLTTLRALGAQKEAISPPDAASAAPLRTITADTFRRLLFRVYSNFLRGKL 180
 Db 121 HDKAVSGLSLTTLRALGAQKEAISPPDAASAAPLRTITADTFRRLLFRVYSNFLRGKL 180
 QY 181 KLYTGEACRTGDR 193
 Db 181 KLYTGEACRTGDR 193

RESULT 4

AAP70256
 ID AAP70256 standard; protein; 193 AA.

AC AAP70256;

DT 19-FEB-1991 (first entry)

DE Sequence of human erythropoietin (EPO).

KW Renal anaemia therapy; hormone.

OS Homo sapiens.

PH Key Location/Qualifiers

FT Peptide 1..27 /label=SIGNALL

FT Protein 28..193

FT Region 81..97

EP232034-A.

PD 12-AUG-1987.

PF 19-JAN-1987; 87EP-0300399.

PR 23-JAN-1986; 86JP-0012868.

PA (SUMO) SUMITOMO CHEM IND KK.

PA (SUMI-) SUMITOMI SEIYAKU KK.

PI Yanagi H, Ogawa I, Okamoto M, Hozumi T, Soga A, Yoshima T;

PI Tsutsumi M;

DR WPI; 1987-223006/32.

DR N-PSDB; AAN70360, AAN70361.

PT Human erythropoietin prodn. - by culturing human cells, esp.

PT Namatawa cells, transformed with DNA encoding human erythropoietin

PS Disclosure; Fig 1; 22pp; English.

XX A cDNA library was prepd. from the poly (A) RNA, which was isolated

CC from the erythropoietin-producing human hepatoma cell Hp-1. The cDNA

CC library was screened using the probes given in AAN70361 and AAN70362. A

CC plasmid (named as p58-A20) was isolated. The nucleotide sequence of

CC the cDNA obtained from this clone is shown in AAN70360.

XX Sequence 193 AA;

Query Match 99.3%; Score 991; DB 8; Length 193;

Best Local Similarity 99.5%; Pred. No. 2.8e-99;

Matches 192; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVECPAMWMLLSTLSPGLPVIGAPPRLLICDSRVLEERYLLLEAKENITTCGAHC 60

Db 1 MGVECPAMWMLLSTLSPGLPVIGAPPRLLICDSRVLEERYLLLEAKENITTCGAHC 60
 QY SLNENITVPDTKVFYAMKREMEVROQAVEWOGIALISEAVLRQALLVNSOPWEPLQL 120
 Db 61 SLNENITVPDTKVFYAMKREMEVROQAVEWOGIALISEAVLRQALLVNSOPWEPLQL 120
 QY 121 HDKAVSGLSLTTLRALGAQKEAISPPDAASAAPLRTITADTFRRLLFRVYSNFLRGKL 180
 Db 121 HDKAVSGLSLTTLRALGAQKEAISPPDAASAAPLRTITADTFRRLLFRVYSNFLRGKL 180
 QY 181 KLYTGEACRTGDR 193
 Db 181 KLYTGEACRTGDR 193

RESULT 5

AAR65499
 ID AAR65499 standard; protein; 193 AA.

AC AAR65499;

DT 24-JUN-1995 (first entry)

DE Human prepro-erythropoietin.

KW Erythropoietin; therapeutic; ss.

OS Synthetic.

PH Key Location/Qualifiers

FT Peptide 1..27 /note="leader peptide"

PN W09425055-A.

PD 10-NOV-1994.

PF 29-APR-1994; 94WO-0504755.

PR 29-APR-1993; 93US-0055076.

PA (ABBO) ABBOTT LAB.

PI Devries PJ, Mellovitz BS, Meuth JL, Okasinski GF;

PI Schaefer VC;

DR WPI; 1994-357906/44.

DR N-PSDB; AAQ74760.

PT Erythropoietin analogues - useful for treatment of anaemia and

PT have enhanced erythropoietic effect.

PS Disclosure; Page 38-39; 56pp; English.

XX DNA encoding human prepro-erythropoietin may be ligated into an

CC expression vector for erythropoietin expression in a CHO

CC cell culture. Site-directed mutagenesis may be used in the

CC construction of EPO analogues with improved activity, which may be

CC used in pharmaceutical compositions for inducing erythropoiesis and

CC treating anaemia.

XX Sequence 193 AA;

Query Match 99.3%; Score 991; DB 15; Length 193;

Best Local Similarity 99.5%; Pred. No. 2.8e-99;

Matches 192; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVECPAMWMLLSTLSPGLPVIGAPPRLLICDSRVLEERYLLLEAKENITTCGAHC 60

1 MGVECPAMWMLLSTLSPGLPVIGAPPRLLICDSRVLEERYLLLEAKENITTCGAHC 60

Db 61 SLNENTVPTDKVNFYAMKRMVEYGOAVEWOGALLSEAVLRGQALLVNSSQPMPEIQL 120

QY 121 HDKAVSGLSRLTTLRALGAQKEAISPPDASAAPLRTITADTFPKLFRVYSNFLRGKL 180
 |||||||

Db 121 HDKAVSGLSRLTTLRALGAQKEAISPPDASAAPLRTITADTFPKLFRVYSNFLRGKL 180

QY 181 KLYTGEACRTGDR 193
 |||||||

Db 181 KLYTGEACRTGDR 193

RESULT 6
 AAR81982
 ID AAR81982 standard; protein: 193 AA.

AC AAR81982;
 XX
 XX 27-FEB-1996 (first entry)

DE Human erythropoietin.
 XX
 XX Erythropoietin; sialylation; sialic acid; glycosylation;
 KW reticulocyte; red blood cell; erythrocyte; haematocrit.
 XX
 OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT Peptide 1..27
 FT /label= sig_peptide 51
 FT Modified-site /label= N-glycosylation_site 65
 FT Modified-site /label= N-glycosylation_site 110
 FT Modified-site /label= N-glycosylation_site 153
 FT Modified-site /label= O-glycosylation_site

PM EP668351-A1.
 XX
 XX 23-AUG-1995.
 PF 12-OCT-1990; 90EP-0101849.
 XX
 XX 09-OCT-1990; 90MO-US05758.
 PR 13-OCT-1989; 89US-0421444.
 XX
 XX (AMGE-) AMGEN INC.
 PA
 XX
 PI Byrne TE, Elliott SG;
 DR WPI: 1995-284791/38.
 XX
 XX New human erythropoietin analogues with increased glycosylation -
 PT have increased activity useful for increasing prodn. of
 PT reticulocytes and red blood cells
 PS Disclosure: Fig 5, 31pp; English.
 XX
 XX Human urinary erythropoietin (AAR81982) is a glycoprotein contg. 3 N-
 CC linked and 1 O-linked oligosaccharide chain. Erythropoietin
 CC analogues (AAR81983-87) have been produced in which the number of
 CC glycosylation sites is increased.
 XX
 SQ Sequence 193 AA;

Query Match 99.3%; Score 991; DB 16; Length 193;
 Best Local Similarity 99.5%; Pred. No. 2.8e-99;
 Matches 192; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVHECPAMLWLLLSLSPGLPVLGAPRLICDSRVLERLLEKKEENITTCGAENC 60
 |||||||
 Db 1 MGVHECPAMLWLLLSLSPGLPVLGAPRLICDSRVLERLLEKKEENITTCGAENC 60

QY 61 SLNENTVPTDKVNFYAMKRMVEYGOAVEWOGALLSEAVLRGQALLVNSSQPMPEIQL 120
 |||||||

Db 61 SLNENTVPTDKVNFYAMKRMVEYGOAVEWOGALLSEAVLRGQALLVNSSQPMPEIQL 120

QY 121 HDKAVSGLSRLTTLRALGAQKEAISPPDASAAPLRTITADTFPKLFRVYSNFLRGKL 180
 |||||||

Db 121 HDKAVSGLSRLTTLRALGAQKEAISPPDASAAPLRTITADTFPKLFRVYSNFLRGKL 180

QY 181 KLYTGEACRTGDR 193
 |||||||

Db 181 KLYTGEACRTGDR 193

RESULT 7
 AAR71137
 ID AAR71137 standard; protein: 193 AA.

AC AAR71137;
 XX
 XX 17-OCT-1995 (first entry)

DE Human erythropoietin.
 XX
 XX Human erythropoietin; glycosylation; sialic acid; solubility;
 KW half-life; biological activity; proteolysis resistance; anaemia;
 KW chronic renal failure.
 XX
 OS Homo sapiens.

XX
 FH Key Location/Qualifiers
 FT Peptide 1..27
 FT /label= sig_peptide 51
 FT W09505465-A.
 XX
 XX 23-FEB-1995.
 PD 16-AUG-1994; 94WO-US09257.
 XX
 XX 13-OCT-1989; 89US-0421444.
 PR 17-AUG-1993; 93US-0108016.
 XX
 XX (AMGE-) AMGEN INC.
 PA
 XX
 PI Byrne TE, Elliott SG;
 DR WPI: 1995-098764/13.
 XX
 XX Erythropoietin (EPO) analogues having additional glycosylation
 PT site(s) - to increase sialic acid content, thereby increasing
 PT solubility, serum half-life, biological activity and resistance
 PT to proteolysis.
 PS Disclosure: Pages 80-81, 108pp; English.
 XX
 XX AAR71137 describes the amino acid sequence of human erythropoietin
 CC (EPO), from which the inventions novel human EPO analogues were
 CC derived. The analogues have at least one additional glycosylation
 CC site, this is used to increase the sialic acid content which in
 CC turn increases the solubility, half-life, biological activity and
 CC proteolysis resistance of the protein. The analogues are useful
 CC in claimed compns. for the treatment of chronic renal failure
 CC associated anaemia.
 XX
 SQ Sequence 193 AA;

Query Match 99.3%; Score 991; DB 16; Length 193;
 Best Local Similarity 99.5%; Pred. No. 2.8e-99;
 Matches 192; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVHECPAMLWLLLSLSPGLPVLGAPRLICDSRVLERLLEKKEENITTCGAENC 60
 |||||||


```

Db      1  MGVHECPAMWLMLLSLSLPLGLPYLGAPPRLLICDSRYLERYLEAKAEENITTCACAEHC 60
QY      61  SLNENITVPDTRKVNRYAKRMREVRQOAVEVWQGLALISEAVLRGQALLVNSSQPEPQL 120
      |||
Db      61  SLNENITVPDTRKVNRYAKRMREVRQOAVEVWQGLALISEAVLRGQALLVNSSQPEPQL 120
QY      121  HVDKAVSGLRSLLTTLRALGQKEAISPDDAASAPLRTITADTFRKLFRRYSNPLRGL 180
      |||
Db      121  HVDKAVSGLRSLLTTLRALGQKEAISPDDAASAPLRTITADTFRKLFRRYSNPLRGL 180
QY      181  KLYTGEACRTGDR 193
      |||
Db      181  KLYTGEACRTGDR 193

RESULT 8
AAR74141
ID  AAR74141 standard; Protein; 193 AA.
XX
AC  AAR74141;
XX
DT  30-OCT-1995 (first entry)
XX
DE  Human erythropoietin.
XX
KW  Erythropoietin; anemia; gene therapy; gene transfer; red blood cell;
KM  RBC; erythrocyte; transformation; myoblast; EPO.
XX
OS  Homo sapiens.
XX
XX  WO9513376-A.
XX  18-MAY-1995.
XX
XX  09-NOV-1994; 94WO-US13066.
XX
PR  07-OCT-1994; 94US-0320480.
PR  10-NOV-1993; 93US-0149871.
XX
PA  (AMGE-) AMGEN INC.
XX  (UYSC-) UNIV SOUTHERN CALIFORNIA.
XX
PI  Hamamori Y, Kedes LH, Samal BB;
XX
XX  WPI: 1995-194095/25.
DR  N-PSDB: AA092296.
XX
PT  Gene therapy for treatment of anaemia - and increasing red blood cell
PS  production by transforming red blood cells with the erythropoietin gene
XX
PS  Disclosure; Page 38-40; 51pp; English.
XX
XX  The amino acid sequence encoded by human EPO cDNA is given in AAR74141.
CC  Transfection of target cells, e.g. myoblasts, with EPO cDNA and
CC  implantation into muscle tissue provides increased RBC prodn.
XX
SQ  Sequence 193 AA;

Query Match          99.3%; Score 991; DB 16; Length 193;
Best Local Similarity 99.5%; Pred. No. 2.8e-99;
Matches 192; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      181  KLYTGEACRTGDR 193
      |||
Db      181  KLYTGEACRTGDR 193

RESULT 9
AAR98397
ID  AAR98397 standard; Protein; 193 AA.
XX
AC  AAR98397;
XX
DT  15-SEP-1996 (first entry)
XX
DE  Human erythropoietin.
XX
KW  Erythropoietin; EPO; anaemia; gene therapy; vector;
KM  scaffold attachment region; SAR element; transgenic animal.
XX
XX  Synthetic.
XX
FH  Key Location/Qualifiers
FT  Peptide 1..27
FT  /label= Sig_peptide
FT  Protein 28..193
XX  /label= Mat_protein
XX
XX  WO9619573-A1.
XX
XX  27-JUN-1996.
XX
XX  18-DEC-1995; 95WO-CA00696.
XX
XX  19-DEC-1994; 94US-0358918.
XX
XX  (CANG-) CANGENE CORP.
XX
PI  Delcuve G;
XX
XX  WPI: 1996-309587/31.
DR  N-PSDB: AAT31529.
DR  N-PSDB: AAT31532.
XX
XX  Recombinant DNA molecule expressing mammalian erythropoietin -
PT  useful to transform cell lines, and for gene therapy, e.g. of
PT  anaemia and other red blood cell disorders
XX
PS  Claim 3; Page 58; 84pp; English.
XX
XX  Human erythropoietin (EPO) (AAR98397) functions to promote erythroid
CC  development, to initiate haemoglobin biosynthesis and to stimulate
CC  proliferation of immature erythroid precursors. It can be obtd.
CC  by stable, long-term expression in mammalian cell hosts transfected
CC  with a vector carrying EPO cDNA (AAT31529) or genomic DNA (AAT31532)
CC  operably linked to an expression control sequence and to 5' and 3'
CC  human apolipoprotein scaffold attachment region (SAR) elements (see
CC  also AAT31530-31). Transgenic animals can be produced that express
CC  the recombinant EPO in their milk.
XX
SQ  Sequence 193 AA;

Query Match          99.3%; Score 991; DB 17; Length 193;
Best Local Similarity 99.5%; Pred. No. 2.8e-99;
Matches 192; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 121 HDKAVSGLSLTLRALGAQKEAISPDAASAPLRTTTADTFKRLFRVYSNPLRGKL 180
 DB 121 HDKAVSGLSLTLRALGAQKEAISPDAASAPLRTTTADTFKRLFRVYSNPLRGKL 180
 QY 181 KLYTGEACRTGDR 193
 DB 181 KLYTGEACRTGDR 193
 RESULT 10
 ID AAY94530 standard: protein; 193 AA.
 AC AAY94530:
 DT 28-NOV-2000 (first entry)
 DE Human erythropoietin protein.
 KW Human: erythropoietin; Epo; glycosylation; anaemia;
 KW chronic renal failure; myelosuppressive therapy; cancer;
 KW viral infection; HIV; blood loss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Peptide 1..27
 FT /label= Signal
 FT Protein 28..193
 FT /label= Erythropoietin
 PN WO200024893-A2.
 PD 04-MAY-2000.
 PF 18-OCT-1999: 99WO-US24435.
 PR 23-OCT-1998: 98US-0178292.
 PA (AMGE-) AMGEN INC.
 PI Egrle JC, Elliott SG, Brown JK;
 DR WPI: 2000-350735/30.
 XX Raising and maintaining hematocrit in a mammal by administering an
 PT effective amount of a hyperglycosylated analog of erythropoietin,
 PT useful for treating anemia associated with myelosuppressive therapy or
 PT excessive blood loss -
 PS Disclosure: Figure 1; 63pp; English.
 CC The present sequence is human erythropoietin (Epo). Epo is a
 CC glycoprotein hormone necessary for the maturation of erythroid
 CC progenitor cells into erythrocytes. It has been discovered that
 CC hyperglycosylated Epo has a longer half-life and greater in vivo
 CC activity than recombinant human Epo. Several hyperglycosylated
 CC Epo mutants (AAY94531 to AAY94544) have been made by in vitro
 CC mutagenesis. Hyperglycosylated Epo analogs are useful as they
 CC may be used instead of recombinant Epo to increase and maintain
 CC the level of red blood cells in mammals. The Epo analogs may be
 CC used to treat or prevent anemia associated with chronic renal
 CC failure, myelosuppressive therapy, certain cancers, viral disease
 CC such as HIV and excessive blood loss.
 SQ Sequence 193 AA:
 Query Match 99.3%; Score 991; DB 21; Length 193;
 Best Local Similarity 99.5%; Pred. No. 2.8e-99;
 Matches 192; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MGVHECPAULMLLSLSTPLGLPVGAPRRLICDSRVLERYLLEKAEKENTTGAENIC 60
 DB 1 MGVHECPAULMLLSLSTPLGLPVGAPRRLICDSRVLERYLLEKAEKENTTGAENIC 60

DB 1 MGVHECPAULMLLSLSTPLGLPVGAPRRLICDSRVLERYLLEKAEKENTTGAENIC 60
 QY 61 SLNENITVPDTKVNPFYAMKRMVEYRQOAVEWQGLALLSEAVLRGQALLVNSSQPMWPIQL 120
 DB 61 SLNENITVPDTKVNPFYAMKRMVEYRQOAVEWQGLALLSEAVLRGQALLVNSSQPMWPIQL 120
 QY 121 HDKAVSGLSLTLRALGAQKEAISPDAASAPLRTTTADTFKRLFRVYSNPLRGKL 180
 DB 121 HDKAVSGLSLTLRALGAQKEAISPDAASAPLRTTTADTFKRLFRVYSNPLRGKL 180
 QY 181 KLYTGEACRTGDR 193
 DB 181 KLYTGEACRTGDR 193
 RESULT 11
 ID AAY93638 standard: protein; 193 AA.
 AC AAY93638:
 DT 25-SEP-2000 (first entry)
 DE Amino acid sequence of a human erythropoietin polypeptide.
 KW Human: erythropoietin; Epo; inhibitor; nuclear factor-kappaB; NF-kappaB;
 KW multi-drug resistance gene; malignant hemopathy; solid tumour;
 KW malignant blood disease; leukaemia; lymphoma; solid cancer.
 OS Homo sapiens.
 PN WO200030587-A2.
 PD 02-JUN-2000.
 PF 24-NOV-1999: 99WO-FR02897.
 PR 25-NOV-1998: 98FR-0014858.
 PA (CNRS) CENT NAT RECH SCI.
 PI Hirsch F, Haeflner A;
 DR WPI: 2000-399901/34.
 DR N-PSDB: AAA46697.
 XX Treatment of haematological or solid tumours using an inhibitor of the
 PT activation of nuclear factor-kappaB, particularly to prevent
 PT development of resistance to chemotherapeutics -
 PS Claim 11; Page 30; 30pp; French.
 CC The present sequence represents a human erythropoietin (EPO) polypeptide.
 CC The human growth hormone protein is used as an inhibitor of the
 CC activation of nuclear factor-kappaB (NF-kappaB). The inhibitor inhibits
 CC activation of NF-kappaB, and thus transcription of the multi-drug
 CC resistance gene (which contains binding sites for NF-kappaB within its
 CC regulatory regions). The inhibitors are used to produce pharmaceuticals
 CC which may be used in the treatment of malignant hemopathy or solid
 CC tumours. The inhibitors are especially used to treat malignant blood
 CC diseases (leukaemia, lymphoma) and solid cancers (of breast or ovary).
 SQ Sequence 193 AA:
 Query Match 99.3%; Score 991; DB 21; Length 193;
 Best Local Similarity 99.5%; Pred. No. 2.8e-99;
 Matches 192; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MGVHECPAULMLLSLSTPLGLPVGAPRRLICDSRVLERYLLEKAEKENTTGAENIC 60
 DB 1 MGVHECPAULMLLSLSTPLGLPVGAPRRLICDSRVLERYLLEKAEKENTTGAENIC 60
 QY 61 SLNENITVPDTKVNPFYAMKRMVEYRQOAVEWQGLALLSEAVLRGQALLVNSSQPMWPIQL 120

```

Db      61 SLNENITVPPTKVFYAMKREVEVGQAAVEWQGLALLSEAVLRQOALLVNSSQWPEPLQL 120
      121 HVDAVSGLSRLTTLRLALGAKQEAISPPDASAPLRTITADFERKLFRRYSNFLRGKL 180
      121 HVDAVSGLSRLTTLRLALGAKQEAISPPDASAPLRTITADFERKLFRRYSNFLRGKL 180
Qy      181 KLYTGEACRTGDR 193
      181 KLYTGEACRTGDR 193
Db      181 KLYTGEACRTGDR 193

RESULT 12
AAV99704
ID      AAV99704 standard; protein: 193 AA.
XX
AC      AAV99704;
XX
DT      15-SEP-2000 (first entry)
XX
DE      Human non-glycosylated erythropoietin NGE.
XX
KM      Human: non-glycosylated erythropoietin; NGE: haematocrit;
KW      antihaemic; anaemia; erythropoiesis promoter.
XX
OS      Homo sapiens.
XX
PN      MO200032772-A2.
XX
PD      08-JUN-2000.
XX
PF      23-NOV-1999; 99WO-US27801.
XX
PR      30-NOV-1998; 98US-0110289.
XX
PA      (PLIL ) LILLY & CO ELI.
XX
PI      Beals JM, Glaesner W, Micanovic R, Milligan RL, Witcher DR;
XX
DR      WPI: 2000-412320/35.
XX
PT      Non-glycosylated erythropoietic compound useful for increasing
XX
PT      haematocrit level in mammal with insufficient haematocrit levels in
XX
PT      conditions such as anemia, comprises protein covalently bonded to
XX
PT      polymer -
XX
PS      Claim 1: Page 91-92; 94pp: English.
XX
CC      The present sequence is the non-glycosylated erythropoietin NGE.
CC
CC      The protein promotes erythropoiesis and can therefore be used to
CC      increase haematocrit levels in mammals with conditions such as
CC      anaemia, in which levels of haematocrit are insufficient. Mutants
CC      derived from the present protein can also be used to treat such
CC      conditions. The analogues, designated NGEAS, do not themselves cause a
CC      significant increase in haematocrit but they acquire that property once
CC      they are derivatised with polyethylene glycol polymers. The analogues
CC      can be produced using a linkerless aldehyde modification process. They
CC      show stability and bioactivity in vivo. The compounds can be
CC      produced by recombinant DNA technology or by chemical procedures such as
CC      solution or solid-phase peptide synthesis.
XX
SQ      Sequence 193 AA;

Query Match          99.3%; Score 991; DB 21: Length 193;
Best local similarity 99.5%; Pred. No. 2,8e-99;
Matches 192; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Db      61 SLNENITVPPTKVFYAMKREVEVGQAAVEWQGLALLSEAVLRQOALLVNSSQWPEPLQL 120
Qy      121 HVDAVSGLSRLTTLRLALGAKQEAISPPDASAPLRTITADFERKLFRRYSNFLRGKL 180
      121 HVDAVSGLSRLTTLRLALGAKQEAISPPDASAPLRTITADFERKLFRRYSNFLRGKL 180
Db      121 HVDAVSGLSRLTTLRLALGAKQEAISPPDASAPLRTITADFERKLFRRYSNFLRGKL 180
Qy      181 KLYTGEACRTGDR 193
      181 KLYTGEACRTGDR 193
Db      181 KLYTGEACRTGDR 193

RESULT 13
AAV43398
ID      AAV43398 standard; protein: 193 AA.
XX
AC      AAV43398;
XX
DT      28-JAN-2000 (first entry)
XX
DE      Human erythropoietin protein sequence.
XX
KM      SAR element: scaffold attachment region: human: apolipoprotein B: tPA:
KW      tissue plasminogen activator; protein expression; gene therapy; lysis;
KW      occlusive coronary artery thrombi; transmural myocardial infarction;
KW      ventricular function; congestive heart failure; acute ischaemic stroke;
KW      acute massive pulmonary embolism; venous thrombosis; arterial thrombosis;
KW      embolism; arteriovenous cannulae occlusion; plasminogen activator;
KW      intravenous catheter clearance; blood clot; erythropoietin.
XX
OS      Homo sapiens.
XX
PN      US5985607-A.
XX
PD      16-NOV-1999.
XX
PF      27-JUN-1997; 97US-0883795.
XX
PR      19-DEC-1994; 94US-0358918.
XX
PA      (CANG-) CANGENE CORP.
XX
PI      Awang G, Delcuve G;
XX
DR      WPI: 2000-012788/01.
XX
DR      N-PSDB: AAZ37201.
XX
PT      Recombinant DNA molecules encoding tissue plasminogen activator
XX
PT      proteins, operatively linked to a scaffold attachment region, useful
XX
PT      for the production of tissue plasminogen activator both in vivo and in
XX
PT      vitro -
XX
PS      Example 2: Fig 3; 49pp: English.
XX
CC      This sequence represents the human erythropoietin protein.
CC
CC      The invention relates to a recombinant DNA molecule adapted for
CC      expression of tissue plasminogen activator (tPA). The DNA molecule
CC      comprise a sequence encoding tPA, an expression control sequence
CC      operatively linked to the tPA sequence, and at least one human
CC      apolipoprotein B scaffold attachment region (SAR) element (the SAR is not
CC      a 5' proximal apolipoprotein B SAR). The SAR element is used to increase
CC      the expression of the coding sequences. The recombinant nucleic acids may
CC      be used for the recombinant production of tPA both in vitro or in vivo
CC      (e.g. as part of a gene therapy procedure). tPA may be administered to
CC      treat and remove blood clots. It is especially useful for the lysis of
CC      occlusive coronary artery thrombi associated with evolving transmural
CC      myocardial infarction to improve ventricular function and reduce the risk
CC      of congestive heart failure. Additionally, it may be used in the
CC      management of acute massive pulmonary embolism, venous thrombosis and
CC      acute ischaemic stroke. Finally, tPA may be used in treating arterial
CC      thrombosis or embolism, arteriovenous cannulae occlusion and intravenous
CC      catheter clearance. In contrast to other plasminogen activators
CC      (e.g. urokinase and streptokinase), the activity of tPA is relatively
CC      localised and (in theory) is less likely to produce systemic haemorrhagic

```

CC disorders.
XX Sequence 193 AA;
SQ Query Match 99.3%; Score 991; DB 21; Length 193;
Best Local Similarity 99.5%; Pred. No. 2.8e-99;
Matches 192; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVECPAMIMLLSLISPLGLPVLGAPPRILICDSRVLEKKEAENITTGCAEHC 60
DB 1 MGVECPAMIMLLSLISPLGLPVLGAPPRILICDSRVLEKKEAENITTGCAEHC 60
QY 61 SLNENITVPDTKYNFYAMKMEYRQQAWEVWQGLALISEAVLRGQALLVNSSQPMPELQL 120
DB 61 SLNENITVPDTKYNFYAMKMEYRQQAWEVWQGLALISEAVLRGQALLVNSSQPMPELQL 120
QY 121 HDKAVSGLRSLTTLRALGAOKEAISPPDAASAPLRTTTADTFKRLFRVYSNPLRGKL 180
DB 121 HDKAVSGLRSLTTLRALGAOKEAISPPDAASAPLRTTTADTFKRLFRVYSNPLRGKL 180
QY 181 KLYTGEACRTGDR 193
DB 181 KLYTGEACRTGDR 193

RESULT 14
AAB85573
ID AAB85573 standard; Protein: 193 AA.
XX AAB85573;
AC AAB85573;
XX 29-OCT-2001 (first entry)
DT 29-OCT-2001 (first entry)
XX Human erythropoietin (EPO) sequence.
DE Human erythropoietin (EPO) sequence.
XX Transgenic; pig; human; erythropoietin; EPO; milk; PMSG; hCG;
KM chorionic gonadotrophic hormone; WAP promoter.
XX Homo sapiens.
OS Homo sapiens.
XX Key Location/Qualifiers
FH Peptide 1..27
FT /note="signal peptide"
FT Protein 28..193
FT /note="mature protein"
XX WO200159074-A1.
PN 16-AUG-2001.
PD 16-AUG-2001.
XX 28-JUN-2000; 2000WO-KR00675.
PF 28-JUN-2000; 2000WO-KR00675.
XX 14-FEB-2000; 2000KR-0006888.
PR 14-FEB-2000; 2000KR-0006888.
XX (KORE-) REPUBLIC KOREA.
PA Chang W, Park J, Seong H, Min K, Yang B, Im G, Lee Y, Lee C;
PI Kim J;
PI WPI: 2001-514656/56.
DR N-PSDB: AAH46972.
XX WPI: 2001-514656/56.
DR N-PSDB: AAH46972.
XX Producing transgenic porcine that secretes human erythropoietin (hEPO)
PT in milk. By introducing vector comprising hEPO genome into fertilized
PT eggs of porcine to which PMSG and hCG were administered, and developing
PT progeny -
XX Claim 4; Fig 3; 21pp; English.
XX The invention relates to producing transgenic pigs (P) that secrete
CC human erythropoietin (hEPO) in milk. The method involves administering
CC PMSG and human chorionic gonadotrophic hormone (hCG) into (P), collecting
CC fertilized eggs after mating, injecting expression vector containing a

CC 2.6 kb WAP promoter, hEPO genome and SV40 poly A DNA into male pronuclei,
CC transplanting them in surrogate mother pig and allowing it to give birth.
CC The method provides transgenic porcine capable of secreting hEPO in their
CC milk, thus producing the expensive useful medicine at a low cost with
CC stability on a large scale, giving a contribution to the improvement of
CC human health. The present sequence represents a human EPO sequence
CC incorporated into the genome of porcine.
XX
SQ Sequence 193 AA;
Query Match 99.3%; Score 991; DB 22; Length 193;
Best Local Similarity 99.5%; Pred. No. 2.8e-99;
Matches 192; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVECPAMIMLLSLISPLGLPVLGAPPRILICDSRVLEKKEAENITTGCAEHC 60
DB 1 MGVECPAMIMLLSLISPLGLPVLGAPPRILICDSRVLEKKEAENITTGCAEHC 60
QY 61 SLNENITVPDTKYNFYAMKMEYRQQAWEVWQGLALISEAVLRGQALLVNSSQPMPELQL 120
DB 61 SLNENITVPDTKYNFYAMKMEYRQQAWEVWQGLALISEAVLRGQALLVNSSQPMPELQL 120
QY 121 HDKAVSGLRSLTTLRALGAOKEAISPPDAASAPLRTTTADTFKRLFRVYSNPLRGKL 180
DB 121 HDKAVSGLRSLTTLRALGAOKEAISPPDAASAPLRTTTADTFKRLFRVYSNPLRGKL 180
QY 181 KLYTGEACRTGDR 193
DB 181 KLYTGEACRTGDR 193

RESULT 15
AAB34978
ID AAB34978 standard; Protein: 193 AA.
XX AAB34978;
AC AAB34978;
XX 27-MAR-2001 (first entry)
DT 27-MAR-2001 (first entry)
XX Human erythropoietin SEQ ID NO: 4.
DE Human erythropoietin SEQ ID NO: 4.
XX Chimpanzee; erythropoietin; EPO; hybridisation probe; gene therapy;
KM mapping; therapeutic agent.
XX Homo sapiens.
OS Homo sapiens.
XX WO200068376-A1.
PN 16-NOV-2000.
PD 16-NOV-2000.
XX 05-MAY-2000; 2000WO-US12370.
PF 05-MAY-2000; 2000WO-US12370.
XX 07-MAY-1999; 99US-0307307.
PR 07-MAY-1999; 99US-0307307.
PR 28-MAR-2000; 2000US-0307307.
XX (GETH) GENENTECH INC.
PA Desauvage F, Henner DJ;
PI Desauvage F, Henner DJ;
PI WPI: 2001-007393/01.
DR WPI: 2001-007393/01.
XX Nucleic acids encoding chimpanzee erythropoietin, useful for treatment
PT of e.g. anemia, also derived proteins, antibodies and modulators -
PT Disclosure; Fig 3; 109pp; English.
XX The present invention provides the coding and protein sequences of
CC chimpanzee erythropoietin (EPO). These sequences can be used in gene
CC therapy, to block the activity of EPO, as hybridisation probes, in
CC genetic and chromosome mapping and as therapeutic agents.
XX Sequence 193 AA;

PS Claim 34; Page 94; 109pp; English.

CC The present invention provides the coding and protein sequences of
CC chimpanzee erythropoietin (EPO). These sequences can be used in gene
CC therapy, to block the activity of EPO, as hybridisation probes, in
CC genetic and chromosome mapping and as therapeutic agents.

XX
SQ Sequence 166 AA;

Query Match 99.5%; Score 839; DB 22; Length 166;
Best Local Similarity 100.0%; Pred. No. 2e-86;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APPRLICDSRVLERYLLEKKEAENITTCGAHCSLNENITVPDTKVNRYAMKRNKXSOQA 60
DB 1 APPRLICDSRVLERYLLEKKEAENITTCGAHCSLNENITVPDTKVNRYAMKRNKXSOQA 60
QY 61 VEWOGIALLSAVALRGQALLVNSSQWPEPLQHDKAVSGRSITTLRALGAQKEAIS 120
DB 61 VEWOGIALLSAVALRGQALLVNSSQWPEPLQHDKAVSGRSITTLRALGAQKEAIS 120
QY 121 PPDAASAPLRTITADTFERKLFRRVYSNPLRGKLYTGECRTGDR 166
DB 121 PPDAASAPLRTITADTFERKLFRRVYSNPLRGKLYTGECRTGDR 166

RESULT 2
AAB35002
ID AAB35002 standard; Protein; 193 AA.
XX AAB35002;
AC
XX
XX 27-MAR-2001 (first entry)
DE Chimpanzee erythropoietin fragment SEQ ID NO: 34.
XX
XX Chimpanzee; erythropoietin; EPO; hybridisation probe; gene therapy;
KM mapping; therapeutic agent.
XX
XX Pan sp.
OS
XX
XX WO200068376-A1.
PN
XX
XX 16-NOV-2000.
PD
XX
XX 05-MAY-2000; 2000WO-US12370.
PE
XX
XX 07-MAY-1999; 99US-0307307.
PR
XX 28-MAR-2000; 2000US-0307307.
PK
XX
XX (GETH) GENENTECH INC.
PA
XX
XX Desauvage F, Hennen DJ;
PI
XX
XX WPI; 2001-007393/01.
DR
XX
XX Nucleic acids encoding chimpanzee erythropoietin, useful for treatment
PT of e.g. anemia, also derived proteins, antibodies and modulators -
XX
XX
XX Claim 35; Page 96; 109pp; English.
PS
XX
XX The present invention provides the coding and protein sequences of
CC chimpanzee erythropoietin (EPO). These sequences can be used in gene
CC therapy, to block the activity of EPO, as hybridisation probes, in
CC genetic and chromosome mapping and as therapeutic agents.

XX
SQ Sequence 193 AA;

Query Match 99.5%; Score 839; DB 22; Length 193;
Best Local Similarity 100.0%; Pred. No. 2.5e-86;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 APPRLICDSRVLERYLLEKKEAENITTCGAHCSLNENITVPDTKVNRYAMKRNKXSOQA 60

DB 28 APPRLICDSRVLERYLLEKKEAENITTCGAHCSLNENITVPDTKVNRYAMKRNKXSOQA 87
QY 61 VEWOGIALLSAVALRGQALLVNSSQWPEPLQHDKAVSGRSITTLRALGAQKEAIS 120
DB 88 VEWOGIALLSAVALRGQALLVNSSQWPEPLQHDKAVSGRSITTLRALGAQKEAIS 147
QY 121 PPDAASAPLRTITADTFERKLFRRVYSNPLRGKLYTGECRTGDR 166
DB 148 PPDAASAPLRTITADTFERKLFRRVYSNPLRGKLYTGECRTGDR 193

RESULT 3
AAB34988
ID AAB34988 standard; Protein; 166 AA.
XX AAB34988;
AC
XX
XX 27-MAR-2001 (first entry)
DE Chimpanzee erythropoietin fragment SEQ ID NO: 20.
XX
XX Chimpanzee; erythropoietin; EPO; hybridisation probe; gene therapy;
KM mapping; therapeutic agent.
XX
XX Pan sp.
OS
XX
XX WO200068376-A1.
PN
XX
XX 16-NOV-2000.
PD
XX
XX 05-MAY-2000; 2000WO-US12370.
PE
XX
XX 07-MAY-1999; 99US-0307307.
PR
XX 28-MAR-2000; 2000US-0307307.
PK
XX
XX (GETH) GENENTECH INC.
PA
XX
XX Desauvage F, Hennen DJ;
PI
XX
XX WPI; 2001-007393/01.
DR
XX
XX Nucleic acids encoding chimpanzee erythropoietin, useful for treatment
PT of e.g. anemia, also derived proteins, antibodies and modulators -
XX
XX
XX Claim 34; Page 94; 109pp; English.
PS
XX
XX The present invention provides the coding and protein sequences of
CC chimpanzee erythropoietin (EPO). These sequences can be used in gene
CC therapy, to block the activity of EPO, as hybridisation probes, in
CC genetic and chromosome mapping and as therapeutic agents.

XX
SQ Sequence 166 AA;

Query Match 99.2%; Score 836; DB 22; Length 166;
Best Local Similarity 99.4%; Pred. No. 4.4e-86;
Matches 165; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 APPRLICDSRVLERYLLEKKEAENITTCGAHCSLNENITVPDTKVNRYAMKRNKXSOQA 60
DB 1 APPRLICDSRVLERYLLEKKEAENITTCGAHCSLNENITVPDTKVNRYAMKRNKXSOQA 60
QY 61 VEWOGIALLSAVALRGQALLVNSSQWPEPLQHDKAVSGRSITTLRALGAQKEAIS 120
DB 61 VEWOGIALLSAVALRGQALLVNSSQWPEPLQHDKAVSGRSITTLRALGAQKEAIS 120
QY 121 PPDAASAPLRTITADTFERKLFRRVYSNPLRGKLYTGECRTGDR 166
DB 121 PPDAASAPLRTITADTFERKLFRRVYSNPLRGKLYTGECRTGDR 166

RESULT 4
AAB35004


```

ID  AAB35004 standard. Protein: 193 AA.
XX
XX  AAB35004;
AC
XX  27-MAR-2001 (first entry)
DT
XX  Chimpanzee erythropoietin fragment SEQ ID NO: 36.
DE
XX  Chimpanzee erythropoietin; EPO; hybridisation probe; gene therapy;
KM  mapping; therapeutic agent.
XX
OS  Pan sp.
XX
XX  WO200068376-A1.
PN
XX  16-NOV-2000.
PD
XX  05-MAY-2000; 2000MO-US12370.
PE
XX  07-MAY-1999; 99US-0307307.
PR  28-MAR-2000; 2000US-0307307.
XX
XX  (GETH ) GENENTECH INC.
PA
XX  Desauvage F, Henner DJ;
PI
XX  WPI; 2001-007393/01.
DR
XX  Nucleic acids encoding chimpanzee erythropoietin, useful for treatment
PT  of e.g. anemia, also derived proteins, antibodies and modulators -
XX
XX  Claim 35; Page 96; 109pp; English.
PS
XX
XX  The present invention provides the coding and protein sequences of
CC  chimpanzee erythropoietin (EPO). These sequences can be used in gene
CC  therapy, to block the activity of EPO, as hybridisation probes, in
CC  genetic and chromosome mapping and as therapeutic agents.
XX
SQ  Sequence 193 AA;

Query Match          99.2%; Score 836; DB 22; Length 193;
Best Local Similarity 99.4%; Pred. No. 5.4e-86;
Matches 165; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY  1 APPRLICDSRVLEERYLLEAKAEENITTCGAHCSLNNITVPDTKYNFYAMKRNXSXQA 60
    |||||||
DB  28 APPRLICDSRVLEERYLLEAKAEENITTCGAHCSLNNITVPDTKYNFYAMKRNXTXQA 87
OY  61 VEWOGIALISEAVLRGQALLVNSQWPPELQLVHDKAVSGLSLTTLLRALGAQKEAIS 120
    |||||||
DB  88 VEWOGIALISEAVLRGQALLVNSQWPPELQLVHDKAVSGLSLTTLLRALGAQKEAIS 147
OY  121 PPDAASAAPLRTITADTFKRLFRVYSNFLRGKILKLYTGECACRTGDR 166
    |||||||
DB  148 PPDAASAAPLRTITADTFKRLFRVYSNFLRGKILKLYTGECACRTGDR 193

RESULT 5
AAB34987
ID  AAB34987 standard. Protein: 166 AA.
XX
XX  AAB34987;
AC
XX  27-MAR-2001 (first entry)
DT
XX  Chimpanzee erythropoietin fragment SEQ ID NO: 19.
DE
XX  Chimpanzee erythropoietin; EPO; hybridisation probe; gene therapy;
KM  mapping; therapeutic agent.
XX
XX  Pan sp.
OS
XX  WO200068376-A1.
PN

```

```

XX  16-NOV-2000.
PD
XX  05-MAY-2000; 2000MO-US12370.
PE
XX  07-MAY-1999; 99US-0307307.
PR  28-MAR-2000; 2000US-0307307.
XX
XX  (GETH ) GENENTECH INC.
PA
XX  Desauvage F, Henner DJ;
PI
XX  WPI; 2001-007393/01.
DR
XX  Nucleic acids encoding chimpanzee erythropoietin, useful for treatment
PT  of e.g. anemia, also derived proteins, antibodies and modulators -
XX
XX  Claim 34; Page 94; 109pp; English.
PS
XX
XX  The present invention provides the coding and protein sequences of
CC  chimpanzee erythropoietin (EPO). These sequences can be used in gene
CC  therapy, to block the activity of EPO, as hybridisation probes, in
CC  genetic and chromosome mapping and as therapeutic agents.
XX
SQ  Sequence 166 AA;

Query Match          99.1%; Score 835; DB 22; Length 166;
Best Local Similarity 99.4%; Pred. No. 5.7e-86;
Matches 165; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY  1 APPRLICDSRVLEERYLLEAKAEENITTCGAHCSLNNITVPDTKYNFYAMKRNXSXQA 60
    |||||||
DB  1 APPRLICDSRVLEERYLLEAKAEENITTCGAHCSLNNITVPDTKYNFYAMKRNXSXQA 60
OY  61 VEWOGIALISEAVLRGQALLVNSQWPPELQLVHDKAVSGLSLTTLLRALGAQKEAIS 120
    |||||||
DB  61 VEWOGIALISEAVLRGQALLVNSQWPPELQLVHDKAVSGLSLTTLLRALGAQKEAIS 120
OY  121 PPDAASAAPLRTITADTFKRLFRVYSNFLRGKILKLYTGECACRTGDR 166
    |||||||
DB  121 PPDAASAAPLRTITADTFKRLFRVYSNFLRGKILKLYTGECACRTGDR 166

RESULT 6
AAB35003
ID  AAB35003 standard. Protein: 193 AA.
XX
XX  AAB35003;
AC
XX  27-MAR-2001 (first entry)
DT
XX  Chimpanzee erythropoietin fragment SEQ ID NO: 35.
DE
XX  Chimpanzee erythropoietin; EPO; hybridisation probe; gene therapy;
KM  mapping; therapeutic agent.
XX
XX  Pan sp.
OS
XX  WO200068376-A1.
PN
XX  16-NOV-2000.
PD
XX  05-MAY-2000; 2000MO-US12370.
PE
XX  07-MAY-1999; 99US-0307307.
PR  28-MAR-2000; 2000US-0307307.
XX
XX  (GETH ) GENENTECH INC.
PA
XX  Desauvage F, Henner DJ;
PI
XX  WPI; 2001-007393/01.
DR
XX

```

PT Nucleic acids encoding chimpanzee erythropoietin, useful for treatment of e.g. anemia, also derived proteins, antibodies and modulators -

XX Claim 35; Page 96; 109pp; English.

PS

CC The present invention provides the coding and protein sequences of chimpanzee erythropoietin (EPO). These sequences can be used in gene therapy, to block the activity of EPO, as hybridisation probes, in genetic and chromosome mapping and as therapeutic agents.

CC

XX

SQ Sequence 193 AA;

Query Match 99.1%; Score 835; DB 22; Length 193;
Best Local Similarity 99.4%; Pred. No. 7e-86; Mismatches 0; Indels 0; Gaps 0;

Matches 165; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLELYLLEAKEENITTCGAHCNSLNNITVPDTKVNPFAMKRNKXQQA 60
DB 28 APPRLICDSRVLELYLLEAKEENITTCGAHCNSLNNITVPDTKVNPFAMKRNKXQQA 87
OY 61 VEWOGGLALISEAVLRGALLVNSSQPWEPIQLHVDKAVSGLSRSLTTLRALGAQKEAIS 120
DB 88 VEWOGGLALISEAVLRGALLVNSSQPWEPIQLHVDKAVSGLSRSLTTLRALGAQKEAIS 147
OY 121 PPDASAAPLRTITADTFRKLFRRVSNFLRGKLIKLYTGACRTGDR 166
DB 148 PPDASAAPLRTITADTFRKLFRRVSNFLRGKLIKLYTGACRTGDR 193

RESULT 7
AAB34989
ID AAB34989 standard; Protein: 166 AA.

AC AAB34989;

DT 27-MAR-2001 (first entry)

DE Chimpanzee erythropoietin fragment SEQ ID NO: 21.

XX

KW Chimpanzee; erythropoietin; EPO; hybridisation probe; gene therapy;

XX mapping; therapeutic agent.

OS Pan sp.

XX WO200068376-A1.

PN 16-NOV-2000.

PD 05-MAY-2000; 2000WO-US12370.

XX

PR 07-MAY-1999; 99US-0307307.

PR 28-MAR-2000; 2000US-0307307.

XX

PA (GETH) GENENTECH INC.

XX

PI Desauvage F, Henner DJ;

XX WPI; 2001-007393/01.

DR

PT Nucleic acids encoding chimpanzee erythropoietin, useful for treatment of e.g. anemia, also derived proteins, antibodies and modulators -

XX

PS Claim 34; Page 94; 109pp; English.

XX

CC The present invention provides the coding and protein sequences of chimpanzee erythropoietin (EPO). These sequences can be used in gene therapy, to block the activity of EPO, as hybridisation probes, in genetic and chromosome mapping and as therapeutic agents.

CC

XX

SQ Sequence 166 AA;

Query Match 98.7%; Score 832; DB 22; Length 166;
Best Local Similarity 98.8%; Pred. No. 1.2e-85;

Matches 164; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLELYLLEAKEENITTCGAHCNSLNNITVPDTKVNPFAMKRNKXQQA 60
DB 1 APPRLICDSRVLELYLLEAKEENITTCGAHCNSLNNITVPDTKVNPFAMKRNKXQQA 87
OY 61 VEWOGGLALISEAVLRGALLVNSSQPWEPIQLHVDKAVSGLSRSLTTLRALGAQKEAIS 120
DB 61 VEWOGGLALISEAVLRGALLVNSSQPWEPIQLHVDKAVSGLSRSLTTLRALGAQKEAIS 147
OY 121 PPDASAAPLRTITADTFRKLFRRVSNFLRGKLIKLYTGACRTGDR 166
DB 121 PPDASAAPLRTITADTFRKLFRRVSNFLRGKLIKLYTGACRTGDR 193

RESULT 8
AAB35005
ID AAB35005 standard; Protein: 193 AA.

AC AAB35005;

DT 27-MAR-2001 (first entry)

DE Chimpanzee erythropoietin fragment SEQ ID NO: 37.

XX

KW Chimpanzee; erythropoietin; EPO; hybridisation probe; gene therapy;

XX mapping; therapeutic agent.

OS Pan sp.

XX WO200068376-A1.

PN 16-NOV-2000.

PD 05-MAY-2000; 2000WO-US12370.

XX

PR 07-MAY-1999; 99US-0307307.

PR 28-MAR-2000; 2000US-0307307.

XX

PA (GETH) GENENTECH INC.

XX

PI Desauvage F, Henner DJ;

XX WPI; 2001-007393/01.

DR

PT Nucleic acids encoding chimpanzee erythropoietin, useful for treatment of e.g. anemia, also derived proteins, antibodies and modulators -

XX

PS Claim 35; Page 96; 109pp; English.

XX

CC The present invention provides the coding and protein sequences of chimpanzee erythropoietin (EPO). These sequences can be used in gene therapy, to block the activity of EPO, as hybridisation probes, in genetic and chromosome mapping and as therapeutic agents.

CC

XX

SQ Sequence 193 AA;

Query Match 98.7%; Score 832; DB 22; Length 193;
Best Local Similarity 98.8%; Pred. No. 1.5e-85;

Matches 164; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLELYLLEAKEENITTCGAHCNSLNNITVPDTKVNPFAMKRNKXQQA 60
DB 28 APPRLICDSRVLELYLLEAKEENITTCGAHCNSLNNITVPDTKVNPFAMKRNKXQQA 87
OY 61 VEWOGGLALISEAVLRGALLVNSSQPWEPIQLHVDKAVSGLSRSLTTLRALGAQKEAIS 120
DB 88 VEWOGGLALISEAVLRGALLVNSSQPWEPIQLHVDKAVSGLSRSLTTLRALGAQKEAIS 147
OY 121 PPDASAAPLRTITADTFRKLFRRVSNFLRGKLIKLYTGACRTGDR 166
DB 148 PPDASAAPLRTITADTFRKLFRRVSNFLRGKLIKLYTGACRTGDR 193

CC with chronic renal failure, myelosuppressive therapy, certain cancers,
 CC viral disease such as HIV and excessive blood loss.
 CC Note: The present sequence is not shown in the specification but is
 CC derived from the wild-type erythropoietin shown in Figure 1 (AAV94530).
 XX

SO Sequence 193 AA;

Query Match 98.0%; Score 826; DB 21; Length 193;
 Best Local Similarity 97.6%; Pred. No. 7.3e-85;
 Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLELYLLEAKAEENITTCGAECISINENITVPDTKVNFMKRNKXQQA 60
 DB 28 APPRLICDSRVLELYLLEAKAEENITTCGAECISINENITVPDTKVNFMKRNKXQQA 87
 OY 61 VEVWOGLLALSEAVLRGQALLVNSSQPMPEIQLHVDKAVSGISLITLLRALGAQKEAIS 120
 DB 88 VEVWOGLLALSEAVLRGQALLVNSSQPMPEIQLHVDKAVSGISLITLLRALGAQKEAIS 147
 OY 121 PPDASAAPLRTITADTFKRLFRVYSNPLRGKRLKLYTGEACRTGDR 166
 DB 148 PPDASAAPLRTITADTFKRLFRVYSNPLRGKRLKLYTGEACRTGDR 193

RESULT 11

AA070398
 ID AAP70398 standard; protein; 166 AA.

XX AC AAP70398;

DT 19-FEB-1991 (first entry)

DE Sequence of human erythropoietin (EPO).

KW Mega-karyocyte-platelet growth factor; hormone;
 KW mega-karyocyte colony stimulating factor; therapy;
 KW small acetyl cholinesterase positive cell;
 KW erythrocyte growth effect.

XX OS Homo sapiens.

PN JP62149624-A.

PD 03-JUL-1987.

PF 15-AUG-1986; 86JP-0191542.

PR 13-SEP-1985; 85JP-0203049.

PR 15-AUG-1986; 86JP-0191542.

PA (KAWA)/ KAWAKITA M.

DR WPI: 1987-224837/32.

PT Megakaryocyte-platelet growth factor - contains as active
 PT component human erythropoietin and is used to treat diseases
 PT caused by decrease in platelets
 PS Disclosure: Page 181; 8pp; Japanese.

XX All of the Cys residues in the SQ are labelled "SH". Megakaryocyte-
 CC platelet growth factor contains human EPO as an active principle.
 CC Human EPO has a megakaryocyte colony-stimulating activity and
 CC increases the ratio of small acetyl cholinesterase positive cell
 CC (Schner) which is immature megakaryocyte. Human EPO effects
 CC megakaryocyte-platelet system other than an erythrocyte growth
 CC effect. Megakaryocyte-platelet growth is usable as a remedy for
 CC diseases caused by a platelet decrease.

SO Sequence 166 AA;

Query Match 97.9%; Score 825; DB 8; Length 166;
 Best Local Similarity 97.6%; Pred. No. 7.6e-85;

Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 APPRLICDSRVLELYLLEAKAEENITTCGAECISINENITVPDTKVNFMKRNKXQQA 60
 DB 1 APPRLICDSRVLELYLLEAKAEENITTCGAECISINENITVPDTKVNFMKRNKXQQA 60
 OY 61 VEVWOGLLALSEAVLRGQALLVNSSQPMPEIQLHVDKAVSGISLITLLRALGAQKEAIS 120
 DB 61 VEVWOGLLALSEAVLRGQALLVNSSQPMPEIQLHVDKAVSGISLITLLRALGAQKEAIS 120
 OY 121 PPDASAAPLRTITADTFKRLFRVYSNPLRGKRLKLYTGEACRTGDR 166
 DB 121 PPDASAAPLRTITADTFKRLFRVYSNPLRGKRLKLYTGEACRTGDR 166

RESULT 12

AA023593
 ID AAR23593 standard; Protein; 166 AA.

XX AC AAR23593;

DT 20-OCT-1992 (first entry)

DE Recombinant hematopoietic molecule portion 2.

KW Erythropoietin; EPO; erythrocytes; IL-3; hematopoiesis.

XX OS Homo sapiens.

PN WO9206116-A.

PD 16-APR-1992.

PF 26-SEP-1991; 91WO-US07053.

PR 28-SEP-1990; 90US-0589958.

PA (ORTHO) ORTHO PHARM CORP.

PI Rosen JI;

DR WPI: 1992-150819/18.

PT Recombinant haematopoietic molecules useful in treating
 PT anaemia(s) - comprise IL-3 or GM-CSF, EPO, G-CSF, IL-5 or M-CSF
 PT and has early and later myeloid differentiation activity

PS Disclosure: Page 32; 82pp; English.

XX This protein sequence given comprises the entire amino acid sequence
 CC of human erythropoietin (EPO). EPO leads to the maturation of
 CC erythrocytes and is therefore designated as a late myeloid
 CC differentiation factor (MDF). Within the scope of the invention
 CC hybrid molecules were produced which contain at least a portion of an
 CC early MDF and at least a portion of a late MDF covalently linked. The
 CC EPO sequence given is effective within the scope of the invention in
 CC full or in a truncated version. Amino acids 7-161 act as a
 CC late MDF when recombined with an early MDF eg. IL-3.
 CC These compounds can be used to promote hematopoiesis in a patient.
 CC The bonding of the early and late factors allows a very high conc. of
 CC late MDF at the surface of a cell which the early MDF is bound. It
 CC also allows the early MDA to act more specifically to stimulate only
 CC the desired lineage, thus reducing undesirable effects. These
 CC compounds are useful for treating anaemias of various origins eg. renal
 CC failure and AIDS. It is easier to produce and administer one
 CC recombinant molecule rather than two separate molecules.

SO Sequence 166 AA;

Query Match 97.9%; Score 825; DB 13; Length 166;
 Best Local Similarity 97.6%; Pred. No. 7.6e-85;
 Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy	1	APPRILCSRYLERLLLAKEAENITTTGCAEHCISINENITVPDTKYVNYAMKRNKXSQQA	60		
Db	1	APPRILCSRYLERLLLEAKAEENITTCGAECISINENITVPDTKYVNYAMKRMVEGQQA	60		
Oy	61	VEWOGLLLSAVYRGQALLVNSSQPMPEPQLAHYDKAVSGLSRTTLRLALGAQKEAIS	12		
Db	61	VEWOGLLLSAVYRGQALLVNSSQPMPEPQLAHYDKAVSGLSRTTLRLALGAQKEAIS	12		
Oy	121	PPDASAAPLRITITADTFRKLFRVYVSNFLRGKTLKLTGEACRTGDR	166		
Db	121	PPDASAAPLRITITADTFRKLFRVYVSNFLRGKTLKLTGEACRTGDR	166		
RESULT 13					
AAW77780					
ID	AAW77780	standard; Protein: 166 AA.			
XX	AAW77780;				
AC	AAW77780;				
XX					
XX	24-NOV-1998	(first entry)			
DT					
XX					
DE	Human EPO receptor agonist polypeptide.				
XX					
KM	Haematopoietic receptor agonist; erythropoietin receptor agonist;				
KM	EPO; human; chimeric protein; stem cell expansion; tumour;				
KM	infection; autoimmune disease; haematopoietic disorder; therapy;				
KM	dendritic cell.				
XX					
OS	Homo sapiens.				
XX					
FH	Key	Location/Qualifiers			
FT	Misc-difference	1..6			
FT	/note= "1-6 amino acids of the N-terminus are				
FT	optionally deleted"				
FT	Misc-difference	162..166			
FT	/note= "1-5 amino acids of the C-terminus are				
FT	optionally deleted"				
FT	Misc-difference	23..24			
FT	/note= "possible positions of new C- and N-termini"				
FT	Misc-difference	24..25			
FT	/note= "possible positions of new C- and N-termini"				
FT	Misc-difference	25..26			
FT	/note= "possible positions of new C- and N-termini"				
FT	Misc-difference	26..27			
FT	/note= "possible positions of new C- and N-termini"				
FT	Misc-difference	27..28			
FT	/note= "possible positions of new C- and N-termini"				
FT	Misc-difference	28..29			
FT	/note= "possible positions of new C- and N-termini"				
FT	Misc-difference	29..30			
FT	/note= "possible positions of new C- and N-termini"				
FT	Misc-difference	30..31			
FT	/note= "possible positions of new C- and N-termini"				
FT	Misc-difference	31..32			
FT	/note= "possible positions of new C- and N-termini"				
FT	Misc-difference	32..33			
FT	/note= "possible positions of new C- and N-termini"				
FT	Misc-difference	33..34			
FT	/note= "possible positions of new C- and N-termini"				
FT	Misc-difference	34..35			
FT	/note= "possible positions of new C- and N-termini"				
FT	Misc-difference	35..36			
FT	/note= "possible positions of new C- and N-termini"				
FT	Misc-difference	36..37			
FT	/note= "possible positions of new C- and N-termini"				
FT	Misc-difference	37..38			
FT	/note= "possible positions of new C- and N-termini"				
FT	Misc-difference	38..39			
FT	/note= "possible positions of new C- and N-termini"				
FT	Misc-difference	39..40			
FT	/note= "possible positions of new C- and N-termini"				
FT	Misc-difference	40..41			
FT	/note= "possible positions of new C- and N-termini"				

FT	Misc-difference	41..42	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 42..43	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 43..44	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 44..45	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 45..46	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 46..47	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 47..48	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 48..49	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 49..50	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 50..51	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 51..52	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 52..53	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 53..54	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 54..55	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 55..56	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 56..57	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 57..58	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 77..78	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 78..79	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 79..80	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 81..82	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 82..83	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 84..85	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 85..86	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 86..87	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 87..88	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 88..89	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 108..109	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 109..110	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 110..111	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 111..112	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 112..113	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 113..114	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 114..115	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 115..116	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 116..117	"possible positions of new C- and N-terminal"
FT	Misc-difference	/note/ 117..118	"possible positions of new C- and N-terminal"

CC stability and refold efficiency.

XX Sequence 166 AA;

Query Match 97.9%; Score 825; DB 19; Length 166;
Best Local Similarity 97.6%; Pred. No. 7.6e-85;
Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 APPRLICDSRYLLEAKAEENITTCGAHCSLNENITVPDTKNFYAMKRNXXSXXQA 60
DB 1 APPRLICDSRYLLEAKAEENITTCGAHCSLNENITVPDTKNFYAMKRNXXSXXQA 60
QY 61 VEWQGIALLSEAVLRGQALLVNSSQWPBPLQLHVDKAVSGLRSLTTLRALGAQKEAIS 120
DB 61 VEWQGIALLSEAVLRGQALLVNSSQWPBPLQLHVDKAVSGLRSLTTLRALGAQKEAIS 120
QY 121 PPDASAAPLRTITADTFRKLFVYSNPLRGKCLKLYTGEACRTGDR 166
DB 121 PPDASAAPLRTITADTFRKLFVYSNPLRGKCLKLYTGEACRTGDR 166

RESULT 15

ABB07030
ID ABB07030 standard; Protein; 166 AA.

AC ABB07030;

DT 21-JUN-2002 (first entry)

DE Modified erythropoietin related gene protein sequence.

KW Modified erythropoietin; EPO.

OS Unidentified.

PN KRL45802-B1.

PD 01-AUG-1998.

PF 31-MAY-1994; 94KR-0012082.

PR 31-MAY-1994; 94KR-0012082.

PA (GLDS) LG CHEM CO LTD.

PI Kim C, Song Y, Lee T;

DR WPI: 2000-234250/20.

DR N-PSDB; ABLS0878.

PT MODIFIED ERYTHROPOIETIN GENE AND EXPRESSION VECTORS THEREOF -

PS Disclosure; Page 14; 15pp; Korean.

CC The present invention describes modified erythropoietin (EPO) genes

CC and expression vectors comprising the genes. The present sequence

CC represents a protein sequence from the present invention.

XX Sequence 166 AA;

Query Match 97.9%; Score 825; DB 21; Length 166;

Best Local Similarity 97.6%; Pred. No. 7.6e-85;

Matches 162; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 APPRLICDSRYLLEAKAEENITTCGAHCSLNENITVPDTKNFYAMKRNXXSXXQA 60
DB 1 APPRLICDSRYLLEAKAEENITTCGAHCSLNENITVPDTKNFYAMKRNXXSXXQA 60
QY 61 VEWQGIALLSEAVLRGQALLVNSSQWPBPLQLHVDKAVSGLRSLTTLRALGAQKEAIS 120
DB 61 VEWQGIALLSEAVLRGQALLVNSSQWPBPLQLHVDKAVSGLRSLTTLRALGAQKEAIS 120
QY 121 PPDASAAPLRTITADTFRKLFVYSNPLRGKCLKLYTGEACRTGDR 166

DB 121 PPDASAAPLRTITADTFRKLFVYSNPLRGKCLKLYTGEACRTGDR 166

Search completed: January 8, 2003, 03:41:28
Job time : 46.39 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 02:35:14 ; Search time 26.8802 Seconds
(without alignments)
690.245 Million cell updates/sec

Title: US-09-813-775c-2

Perfect score: 998

Sequence: 1 MGVECPAMWMLLLSLSLP.....NFLRGKIKLYTGACRTGDR 193

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	991	99.3	193	1 ZUHU	erythropoietin pre
2	899.5	90.1	192	1 JG0173	erythropoietin pre
3	894.5	89.6	192	1 I84613	erythropoietin pre
4	794.5	79.6	188	1 I46083	erythropoietin pre
5	778.5	78.0	182	1 S28148	erythropoietin pre
6	773.5	77.1	194	1 I46401	erythropoietin pre
7	769.5	76.0	190	2 I46578	erythropoietin - p
8	758.5	75.0	192	1 A24902	erythropoietin pre
9	731	73.2	195	2 JG7699	erythropoietin - r
10	714.5	71.6	175	2 I46199	erythropoietin - d
11	107	10.7	353	2 G02729	thrombopoietin - h
12	96.5	9.7	353	2 I80105	thrombopoietin pre
13	93.5	9.4	286	2 A55530	megakaryocyte grow
14	90	9.0	356	2 A45330	thrombopoietin - m
15	87.5	8.8	346	2 AE0959	Solute binding rec
16	87	8.7	323	2 AB0333	ribonucleoside-dip
17	86.5	8.7	326	2 UC4135	thrombopoietin pre
18	86.5	8.7	622	2 T02244	probable DNA repl
19	85.5	8.6	897	2 A54696	EGF receptor subst
20	85	8.5	339	2 A83274	UDP-N-acetylpyruvo
21	84.5	8.5	2201	2 AH0095	probable sideroph
22	84	8.4	235	2 JC4227	thrombopoietin pre
23	83.5	8.4	475	2 D84064	succinate-semialde
24	83	8.3	451	2 S75569	hypothetical prote
25	82.5	8.3	296	2 AI0443	probable 2-hydroxy
26	82.5	8.3	480	2 S56639	ribosomal protein
27	82.5	8.3	3033	1 GNMVJ8	genome polypeptide
28	82	8.2	1336	2 T18288	AAC transport prot
29	82	8.2	1980	2 S54307	myosin heavy chain

30	81	8.1	813	2 AF0526	ATP-dependent heli
31	79.5	8.0	929	2 B41863	two-component regu
32	79	7.9	623	2 T03800	anthranilate synth
33	79	7.9	1089	2 S53978	PSI1 protein - yea
34	78.5	7.9	1829	2 T35681	probable sensory h
35	78	7.8	282	2 B37994	KR2 protein - salm
36	78	7.8	743	2 D75590	methyl-accepting c
37	77.5	7.8	242	2 AD1928	hypothetical prote
38	77.5	7.8	896	2 S43074	epidermal growth f
39	77	7.7	304	2 A87258	hypothetical prote
40	77	7.7	637	2 S75772	hypothetical prote
41	76.5	7.7	249	2 G87105	probable amidotran
42	76.5	7.7	567	2 T08405	hypothetical prote
43	76.5	7.7	925	2 E83529	sensor/response re
44	76	7.6	384	2 AI2962	cellulose synthesi
45	76	7.6	389	2 E98320	hypothetical prote

ALIGNMENTS

RESULT 1

ZUHU
erythropoietin precursor [validated] - human
C:Species: Homo sapiens (man)
C:Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 08-Dec-2000
C:Accession: A01855; A24744; A25384; A22210; S56178
R:Jacobs, K.; Shoemaker, C.; Ruderdorf, R.; Neill, S.D.; Kaufman, R.J.; Mifson, A.;
Nature 313, 806-810, 1985
A:Title: Isolation and characterization of genomic and cDNA clones of human erythropo
A:Reference number: A01855; MUID:85137899; PMID:3838366
A:Accession: A01855
A:Molecule type: RNA: DNA
A:Residues: 1-193 <I>
A:Cross-references: GB:X02157; GB:X02158
R:Lin, F.K.; Suggs, S.; Lin, C.H.; Browne, J.K.; Smalling, R.; Egrie, J.C.; Chen, K.K
Proc. Natl. Acad. Sci. U.S.A. 82, 7580-7584, 1985
A:Title: Cloning and expression of the human erythropoietin gene.
A:Reference number: A24744; MUID:86067948; PMID:3865178
A:Accession: A24744
A:Molecule type: DNA
A:Residues: 1-193 <I>
A:Cross-references: GB:M11319; NID:g182197; PIDN:AAA52400.1; PID:g182198
R:Lat, P.H.; Everett, R.; Wang, F.F.; Arakawa, T.; Goldwasser, E.
J. Biol. Chem. 261, 3116-3121, 1986
A:Title: Structural characterization of human erythropoietin.
A:Reference number: A25384; MUID:86140080; PMID:3949763
A:Accession: A25384
A:Molecule type: protein
A:Residues: 28-86, 'Q', 87-193 <I>
A:Experimental source: urine
A:Note: forms without the carboxyl-terminal residue and the four carboxyl-terminal re
R:Yanagawa, S.; Hirade, K.; Ohnoka, H.; Sasaki, R.; Chiba, H.; Ueda, M.; Goto, M.
J. Biol. Chem. 259, 2707-2710, 1984
A:Title: Isolation of human erythropoietin with monoclonal antibodies.
A:Reference number: A22210; MUID:84135751; PMID:6698989
A:Accession: A22210
A:Molecule type: protein
A:Residues: 28-29, 'X', 31-33, 'L', 35-50, 'X', 52-53, 'D', 55, 'G', 57 <I>
R:Matsumoto, S.; Ikura, K.; Ueda, M.; Sasaki, R.
Plant Mol. Biol. 27, 1163-1172, 1995
A:Title: Characterization of a human glycoprotein (erythropoietin) produced in cultur
A:Reference number: S56178; MUID:95284365; PMID:7766897
A:Accession: S56178
A:Molecule type: protein
A:Residues: 28-33, 'X', 35-37 <I>
C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live
C:Genetics:
A:Gene: GDB:EPO
A:Cross-references: GDB:119110; OMIM:133170
A:Map position: 7q21.3-7q22.1
A:Introns: 5/1; 53/3; 82/3; 142/3
C:Function:

A:Description: the primary inducer of erythrocyte formation
C:Superfamily: erythropoietin
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-193/Product: erythropoietin #status experimental <MAT>
F:34-188,56-60/Disulfide bonds: #status experimental
F:51,65,110/Binding site: carbohydrate (Ser) (covalent) #status experimental
F:152/Binding site: carbohydrate (Ser) (covalent) #status experimental

Query Match 99.3%; Score 991; DB 1; Length 193;
Best local Similarity 99.5%; Pred. No. 1,7e-85;
Matches 192; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVECPAMWMLLSLSPGLGPGVPGAPRLICDSRYLERYLEKAEENITTCGAHC 60
|||||
DB 1 MGVECPAMWMLLSLSPGLGPGVPGAPRLICDSRYLERYLEKAEENITTCGAHC 60
|||||
QY 61 SLNENITVPPTKYNFYAMKMEYRQOAVEWQGLALSEAVLNGOALLVNSSQPMPELQ 120
|||||
DB 61 SLNENITVPPTKYNFYAMKMEYRQOAVEWQGLALSEAVLNGOALLVNSSQPMPELQ 120
|||||
QY 121 HDKAVSGLSRLTTLRALGAQKEAISPPDASAPLRTITADTFKRLFRVYSNPLRGKL 180
|||||
DB 121 HDKAVSGLSRLTTLRALGAQKEAISPPDASAPLRTITADTFKRLFRVYSNPLRGKL 180
|||||
QY 181 KLYTGEACRTGDR 193
|||||
DB 181 KLYTGEACRTGDR 193

RESULT 2

JQ0173
erythropoietin precursor - crab-eating macaque
C:Species: Macaca fascicularis (crab-eating macaque)
C>Date: 07-Sep-1990 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999
C:Accession: JQ0173
Gene 44, 201,209, 1986
A:Title: Monkey erythropoietin gene: cloning, expression and comparison with the human
A:Reference number: JQ0173; MUID:87055236; PMID:2877922
A:Accession: JQ0173
A:Molecule type: mRNA
A:Residues: 1-192 <LIN>

A:Cross-references: GB:M18189; GB:M15818; GB:M15819; GB:M18188; NID:9342093; PIDN:AAA368
A:Experimental source: kidney
C:Comment: This protein is the principal hormone involved in the regulation of erythrocy
C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by liver
C:Function:
A:Description: the primary inducer of erythrocyte formation
C:Superfamily: erythropoietin
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-192/Product: erythropoietin #status predicted <MAT>
F:34-187,56-60/Disulfide bonds: #status predicted
F:51,65,110/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:152/Binding site: carbohydrate (Ser) (covalent) #status predicted

Query Match 90.1%; Score 899.5; DB 1; Length 192;
Best local Similarity 91.2%; Pred. No. 6,4e-77;
Matches 176; Conservative 8; Mismatches 8; Indels 1; Gaps 1;

QY 1 MGVECPAMWMLLSLSPGLGPGVPGAPRLICDSRYLERYLEKAEENITTCGAHC 60
|||||
DB 1 MGVECPAMWMLLSLSPGLGPGVPGAPRLICDSRYLERYLEKAEENITTCGAHC 60
|||||
QY 61 SLNENITVPPTKYNFYAMKMEYRQOAVEWQGLALSEAVLNGOALLVNSSQPMPELQ 120
|||||
DB 61 SLNENITVPPTKYNFYAMKMEYRQOAVEWQGLALSEAVLNGOALLVNSSQPMPELQ 120
|||||
QY 121 HDKAVSGLSRLTTLRALGAQKEAISPPDASAPLRTITADTFKRLFRVYSNPLRGKL 180
|||||
DB 121 HDKAVSGLSRLTTLRALGAQKEAISPPDASAPLRTITADTFKRLFRVYSNPLRGKL 179

QY 181 KLYTGEACRTGDR 193
|||||
DB 180 KLYTGEACRTGDR 192

RESULT 3

184613
erythropoietin precursor - rhesus macaque
C:Species: Macaca mulatta (rhesus macaque)
C>Date: 02-Aug-1996 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999
C:Accession: 184613
R:Men, D.; Boissel, J.
Blood 82, 1507-1516, 1993
A:Title: Erythropoietin structure-function relationships: High degree of sequence hom
A:Reference number: 146083; MUID:93372347; PMID:8364201
A:Accession: 184613
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
A:Residues: 1-192 <RES>
A:Cross-references: GB:L10609; NID:9342095; PIDN:AAA36842.1; PID:9342096
C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live
C:Function:
A:Description: the primary inducer of erythrocyte formation
C:Superfamily: erythropoietin
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-192/Product: erythropoietin #status predicted <MAT>
F:34-187,56-60/Disulfide bonds: #status predicted
F:51,65,110/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:152/Binding site: carbohydrate (Ser) (covalent) #status predicted

Query Match 89.6%; Score 894.5; DB 1; Length 192;
Best local Similarity 90.2%; Pred. No. 1,9e-76;
Matches 174; Conservative 10; Mismatches 8; Indels 1; Gaps 1;

QY 1 MGVECPAMWMLLSLSPGLGPGVPGAPRLICDSRYLERYLEKAEENITTCGAHC 60
|||||
DB 1 MGVECPAMWMLLSLSPGLGPGVPGAPRLICDSRYLERYLEKAEENITTCGAHC 60
|||||
QY 61 SLNENITVPPTKYNFYAMKMEYRQOAVEWQGLALSEAVLNGOALLVNSSQPMPELQ 120
|||||
DB 61 SLNENITVPPTKYNFYAMKMEYRQOAVEWQGLALSEAVLNGOALLVNSSQPMPELQ 120
|||||
QY 121 HDKAVSGLSRLTTLRALGAQKEAISPPDASAPLRTITADTFKRLFRVYSNPLRGKL 180
|||||
DB 121 HDKAVSGLSRLTTLRALGAQKEAISPPDASAPLRTITADTFKRLFRVYSNPLRGKL 179
|||||
QY 181 KLYTGEACRTGDR 193
|||||
DB 180 KLYTGEACRTGDR 192

RESULT 4

146083
erythropoietin precursor - cat (fragment)
C:Species: Felis silvestris catus (domestic cat)
C>Date: 16-Aug-1996 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999
C:Accession: 146083
R:Men, D.; Boissel, J.
Blood 82, 1507-1516, 1993
A:Title: Erythropoietin structure-function relationships: High degree of sequence hom
A:Reference number: 146083; MUID:93372347; PMID:8364201
A:Accession: 146083
A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA
A:Residues: 1-188 <WEN>
A:Cross-references: GB:L10606; NID:9163820; PIDN:AAA30807.1; PID:9163821
C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live
C:Function:
A:Description: the primary inducer of erythrocyte formation
C:Superfamily: erythropoietin
C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver
F:1-22/Domain: signal sequence (fragment) #status predicted <SIG>

F:23-188/Product: erythropoietin #status predicted <MAT>
 F:29-183,51-55/Disulfide bonds: #status predicted
 F:46,60,105/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:148/Binding site: carbohydrate (Ser) (covalent) #status predicted

Query Match

79.68; Score 794.5; DB 1; Length 188;

Best Local Similarity 84.18; Pred. No. 4,3e-67;

Matches 159; Conservative 9; Mismatches 20; Indels 1; Gaps 1;

OY 5 ECPAMLLSLSLPLGLPVLGAPPLICDSRYLERYLEAKEENITTCACGSCINE 64
 DB 1 ECPA-LTLLSLSLPLGLPVLGAPPLICDSRYLERYLEAKEENITTCACGSCSE 59
 OY 65 NITVPDTKVFYFAMKREVEVQOAVVEWQGLALLSEAVLRGQALLVNSSQPEPEQLHYDK 124
 DB 60 NITVPDTKVFYFAMKREVEVQOAVVEWQGLALLSEAVLRGQALLVNSSQPEPEQLHYDK 119
 OY 125 AVSGLRSRLTLRLALGAKQKESIPPDAAAPLRTITADTFKRLFRVYSNFLRGKLLYT 184
 DB 120 AVSGLRSRLTLRLALGAKQKESIPPDAAAPLRTITADTFKRLFRVYSNFLRGKLLYT 179
 OY 185 GEACRTGDR 193
 DB 180 GEACRGRDR 188

RESULT 5

S28148
 erythropoietin precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 22-Nov-1993 #sequence_revision 15-Nov-1996 #text_change 21-Jul-2000

C/Accession: S28148; 162743

R:Nagao, M.; Suga, H.; Okano, M.; Masuda, S.; Narita, H.; Ikura, K.; Sasaki, R.

Biochim. Biophys. Acta 1171, 99-102, 1992

A>Title: Nucleotide sequence of rat erythropoietin.

A:Reference number: S28148; MUID:93042015; PMID:1420369

A:Accession: S28148

A:Molecule type: mRNA

A:Residues: 1-192 <NAG>

A:Cross-references: GB:D10763; NID:g220735; PIDN:BA01593.1; PID:g220736

R:Men, D.; Boissel, J.

Blood 82, 1507-1516, 1993

A>Title: Erythropoietin structure-function relationships: High degree of sequence homolo

A:Reference number: I46083; MUID:93372347; PMID:8364201

A:Accession: I62743

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 4-192 <RES>

A:Cross-references: GB:L10608; NID:g204060; PIDN:AAA41126.1; PID:g204061

C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by liver

C:Function: the primary inducer of erythrocyte formation

C:Superfamily: erythropoietin

C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver

F:1-26/Domain: signal sequence #status predicted <SIG>

F:27-192/Product: erythropoietin #status predicted <MAT>

F:33-187,55-165/Disulfide bonds: #status predicted

F:50,64,109/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 78.08; Score 778.5; DB 1; Length 192;

Best Local Similarity 80.88; Pred. No. 1.4e-65;

Matches 156; Conservative 14; Mismatches 22; Indels 1; Gaps 1;

OY 1 MGVECPAMLLSLSLPLGLPVLGAPPLICDSRYLERYLEAKEENITTCACGSCINE 60
 DB 1 MGVEPERT-LTLLSLSLPLGLPVLGAPPLICDSRYLERYLEAKEENITTCACGSCSE 59
 OY 61 SLNENITVPDTKVFYFAMKREVEVQOAVVEWQGLALLSEAVLRGQALLVNSSQPEPEQL 120
 DB 60 SLNENITVPDTKVFYFAMKREVEVQOAVVEWQGLALLSEAVLRGQALLVNSSQPEPEQL 119
 OY 121 HVKAVSGLRSRLTLRLALGAKQKESIPPDAAAPLRTITADTFKRLFRVYSNFLRGKLL 180
 DB 120 HVKAVSGLRSRLTLRLALGAKQKESIPPDAAAPLRTITADTFKRLFRVYSNFLRGKLL 179

DB 120 HIDKAIISGLRSRLTLRLALGAKQKESIPPDAAAPLRTITADTFKRLFRVYSNFLRGKLL 179
 OY 181 KLYTGEACRTGDR 193
 DB 180 KLYTGEACRGRDR 192

RESULT 6

I46401
 erythropoietin precursor - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C>Date: 16-Aug-1996 #sequence_revision 15-Nov-1996 #text_change 22-Jun-1999

C/Accession: I46401; 147077

R:Fu, P.; Evans, B.; Lim, G.B.; Moritz, K.; Wintour, E.M.

Mol. Cell. Endocrinol. 93, 107-116, 1993

A>Title: The sheep erythropoietin gene: molecular cloning and effect of hemorrhage on

A:Reference number: I46401; MUID:93351736; PMID:8349021

A:Accession: I46401

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-194 <FOX>

A:Cross-references: EMBL:Z24681; NID:g395049; PIDN:CA08048.1; PID:g395050

R:Men, D.; Boissel, J.

Blood 82, 1507-1516, 1993

A>Title: Erythropoietin structure-function relationships: High degree of sequence hom

A:Reference number: I46083; MUID:93372347; PMID:8364201

A:Accession: I47077

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 4-15, 'L', 17-107, 'P', 109-194 <WEN>

A:Cross-references: GB:L10610; NID:g165876; PIDN:AAA31518.1; PID:g165877

C:Comment: Erythropoietin is produced by kidney or liver of adult mammals and by live

C:Function: the primary inducer of erythrocyte formation

C:Superfamily: erythropoietin

C:Keywords: erythropoiesis; glycoprotein; hormone; kidney; liver

F:1-27/Domain: signal sequence #status predicted <SIG>

F:28-194/Product: erythropoietin #status predicted <MAT>

F:34-189,56-60/Disulfide bonds: #status predicted

F:51,65,110/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:134/Binding site: carbohydrate (Ser) (covalent) #status predicted

Query Match 77.58; Score 773.5; DB 1; Length 194;

Best Local Similarity 79.44; Pred. No. 4.2e-65;

Matches 154; Conservative 11; Mismatches 28; Indels 1; Gaps 1;

OY 1 MGVECPAMLLSLSLPLGLPVLGAPPLICDSRYLERYLEAKEENITTCACGSCINE 60
 DB 1 MGARDCTPLLLSLPLGLPVLGAPPLICDSRYLERYLEAKEENITTCACGSC 60
 OY 61 SLNENITVPDTKVFYFAMKREVEVQOAVVEWQGLALLSEAVLRGQALLVNSSQPEPEQL 120
 DB 61 SFSENITVPDTKVFYFAMKREVEVQOAVVEWQGLALLSEAVLRGQALLVNSSQPEPEQL 120
 OY 121 HVKAVSGLRSRLTLRLALGAKQKESIPPDAA-SAAPLRTITADTFKRLFRVYSNFLRGK 179
 DB 121 HVKAVSGLRSRLTLRLALGAKQKESIPPDATPSAAPLRTITADTFKRLFRVYSNFLRGK 180
 OY 180 LKTYTGEACRTGDR 193
 DB 181 LTYTGEACRGRDR 194

RESULT 7

I46578
 erythropoietin - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 16-Jul-1999

C/Accession: I46578

R:Men, D.; Boissel, J.

Blood 82, 1507-1516, 1993

A>Title: Erythropoietin structure-function relationships: High degree of sequence hom

A:Reference number: I46083; MUID:93372347; PMID:8364201

0Y	5	BCPAMILLISLSPILGIPVIGAPRLICDSRVLETLERKKAENITGGAHEHSLNE	64
Db	1	BCPA-LLLLSLTLPLGIPVIGAPRLICDSRVLETLERKKAENITGGAQGCSFE	59
0Y	65	NITVDDTKNFPAMKMEVROOAVEWMOGLALISEAVILRGQALLVNSOPWPEIOLHYVK	124
Db	60	NITVDDTKNFPAMKMDVGOALEWMOGLALISEAVILRGQALLVNSQSPSEFPOLHYVK	119
0Y	125	AVSGRLSLTLTLRALGAOKEAISPDDAASAPLRTTADTFRKLFVYVSNPLRGL	180
Db	120	AVSLSRLSLTLRALGAOKEAISPDDAASAPLRTTADTFRKLRITVSNPLRGL	175

```

RESULT 11
G02729
thrombopoietin - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 05-Nov-1999
C:Accession: G02729
R:Im: S.
submitted to the EMBL Data Library, May 1996
A:Reference number: H01637
A:Accession: G02729
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-353 <IMX>
A:Cross-references: EMBL:U59493; NID:g1401245; PIDN:AA803392.1; PID:g1401244
C:Genetics:
A:Gene: hTPO

```

[illegible]

RESULT 12
180105
thrombopoietin precursor - human
N:Alternate names: c-MPL ligand; megakaryocyte growth and development factor precursor
C:Species: Homo sapiens (man)
C:Date: 24-May-1996 #sequence,revision 24-May-1996 #text,change 20-Jun-2000
C:Accession: 159281; 180105; S45331; S48740; I36672; I52610
R:Poster, D.C.; Sprecher, C.A.; Grant, F.J.; Kramet, J.M.; Kuijper, J.L.; Holly, R.D.;
Proc. Natl. Acad. Sci. U.S.A. 91, 13023-13027, 1994
A:Title: Human thrombopoietin: gene structure, cDNA sequence, expression, and chromosomal
A:Reference number: 159281; MUID:95108091; PMID:7809166
A:Accession: 159281
A:Status: translated from GR/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-953 <RES>
A:Cross-references: GB:L36051; NID:g533214; PIDN:AAC37568.1; PID:g533215
A:Accession: 180105
A:Status: translated from GR/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-953 <RES>
A:Cross-references: GB:L36052; NID:g533216; PIDN:AAC37566.1; PID:g533217
R:de Sauvage, F.J.; Hass, P.E.; Spencer, S.D.; Malloy, B.E.; Gurney, A.L.; Spencer, S.A.
D.V.; Eaton, D.L.
Nature 369, 533-538, 1994
A:Title: Stimulation of megakaryocytopoiesis and thrombopoiesis by the c-MPL ligand.

A:Reference number: S45331; MUID:94261202; PMID:8202154
A:Accession: S45331
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-353 <SNU>
A:Cross-references: GB:L33410; NID:g506826; PIDB:AAA59857.1; PID:g506827
R:Solhma, Y.; Akahori, H.; Seki, N.; Hori, T.; Ogami, K.; Kato, T.; Shinada, Y.; Kawam
FERS Lett. 353, 57-61, 1994
A:Title: Molecular cloning and chromosomal localization of the human thrombopoietin g
A:Reference number: S48740; MUID:95010765; PMID:7926023
A:Accession: S48740
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <SOH>
A:Cross-references: GB:D32046; NID:g5773319; PIDB:BAA06807.1; PID:g5773320
R:Batley, T.D.; Bogenberger, J.; Hunt, P.; Li, Y.S.; Lu, H.S.; Martin, F.; Chang, M.
Cell 77, 1117-1124, 1994
A:Title: Identification and cloning of a megakaryocyte growth and development factor
A:Reference number: A54463; MUID:94291201; PMID:8020099
A:Accession: 138672
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-112, 'E', 114-353 <RE3>
A:Cross-references: EMBL:U1025; NID:g511223; PIDB:AAA50553.1; PID:g558078
R:Günery, A.L.; Xiang, W.J.; Xie, M.H.; Malloy, B.E.; Eaton, D.L.; de Sauvage, F.J.
Blood 85, 981-988, 1995
A:Title: Genomic structure, chromosomal localization, and conserved alternative splic
A:Reference number: I52610; MUID:95152076; PMID:7849339
A:Accession: 152610
A:Status: preliminary; translated from GR/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-353 <RE4>
A:Cross-references: GB:S76771; NID:g914225; PIDB:AAB33390.1; PID:g914226
A:Accession: S48740
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <SOH>
A:Cross-references: GB:D32046; NID:g5773319; PIDB:BAA06807.1; PID:g5773320
R:Batley, T.D.; Bogenberger, J.; Hunt, P.; Li, Y.S.; Lu, H.S.; Martin, F.; Chang, M.
Cell 77, 1117-1124, 1994
A:Title: Identification and cloning of a megakaryocyte growth and development factor
A:Reference number: A54463; MUID:94291201; PMID:8020099
A:Accession: 138672
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-112, 'E', 114-353 <RE3>
A:Cross-references: EMBL:U1025; NID:g511223; PIDB:AAA50553.1; PID:g558078
R:Günery, A.L.; Xiang, W.J.; Xie, M.H.; Malloy, B.E.; Eaton, D.L.; de Sauvage, F.J.
Blood 85, 981-988, 1995
A:Title: Genomic structure, chromosomal localization, and conserved alternative splic
A:Reference number: I52610; MUID:95152076; PMID:7849339
A:Accession: 152610
A:Status: preliminary; translated from GR/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-353 <RE4>
A:Cross-references: GB:S76771; NID:g914225; PIDB:AAB33390.1; PID:g914226
A:Accession: S48740
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <SOH>
A:Cross-references: GB:D32046; NID:g5773319; PIDB:BAA06807.1; PID:g5773320
R:Batley, T.D.; Bogenberger, J.; Hunt, P.; Li, Y.S.; Lu, H.S.; Martin, F.; Chang, M.
Cell 77, 1117-1124, 1994
A:Title: Identification and cloning of a megakaryocyte growth and development factor
A:Reference number: A54463; MUID:94291201; PMID:8020099
A:Accession: 138672
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-112, 'E', 114-353 <RE3>
A:Cross-references: EMBL:U1025; NID:g511223; PIDB:AAA50553.1; PID:g558078
R:Günery, A.L.; Xiang, W.J.; Xie, M.H.; Malloy, B.E.; Eaton, D.L.; de Sauvage, F.J.
Blood 85, 981-988, 1995
A:Title: Genomic structure, chromosomal localization, and conserved alternative splic
A:Reference number: I52610; MUID:95152076; PMID:7849339
A:Accession: 152610
A:Status: preliminary; translated from GR/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-353 <RE4>
A:Cross-references: GB:S76771; NID:g914225; PIDB:AAB33390.1; PID:g914226
A:Accession: S48740
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <SOH>
A:Cross-references: GB:D32046; NID:g5773319; PIDB:BAA06807.1; PID:g5773320
R:Batley, T.D.; Bogenberger, J.; Hunt, P.; Li, Y.S.; Lu, H.S.; Martin, F.; Chang, M.
Cell 77, 1117-1124, 1994
A:Title: Identification and cloning of a megakaryocyte growth and development factor
A:Reference number: A54463; MUID:94291201; PMID:8020099
A:Accession: 138672
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-112, 'E', 114-353 <RE3>
A:Cross-references: EMBL:U1025; NID:g511223; PIDB:AAA50553.1; PID:g558078
R:Günery, A.L.; Xiang, W.J.; Xie, M.H.; Malloy, B.E.; Eaton, D.L.; de Sauvage, F.J.
Blood 85, 981-988, 1995
A:Title: Genomic structure, chromosomal localization, and conserved alternative splic
A:Reference number: I52610; MUID:95152076; PMID:7849339
A:Accession: 152610
A:Status: preliminary; translated from GR/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-353 <RE4>
A:Cross-references: GB:S76771; NID:g914225; PIDB:AAB33390.1; PID:g914226
A:Accession: S48740
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <SOH>
A:Cross-references: GB:D32046; NID:g5773319; PIDB:BAA06807.1; PID:g5773320
R:Batley, T.D.; Bogenberger, J.; Hunt, P.; Li, Y.S.; Lu, H.S.; Martin, F.; Chang, M.
Cell 77, 1117-1124, 1994
A:Title: Identification and cloning of a megakaryocyte growth and development factor
A:Reference number: A54463; MUID:94291201; PMID:8020099
A:Accession: 138672
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-112, 'E', 114-353 <RE3>
A:Cross-references: EMBL:U1025; NID:g511223; PIDB:AAA50553.1; PID:g558078
R:Günery, A.L.; Xiang, W.J.; Xie, M.H.; Malloy, B.E.; Eaton, D.L.; de Sauvage, F.J.
Blood 85, 981-988, 1995
A:Title: Genomic structure, chromosomal localization, and conserved alternative splic
A:Reference number: I52610; MUID:95152076; PMID:7849339
A:Accession: 152610
A:Status: preliminary; translated from GR/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-353 <RE4>
A:Cross-references: GB:S76771; NID:g914225; PIDB:AAB33390.1; PID:g914226
A:Accession: S48740
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <SOH>
A:Cross-references: GB:D32046; NID:g5773319; PIDB:BAA06807.1; PID:g5773320
R:Batley, T.D.; Bogenberger, J.; Hunt, P.; Li, Y.S.; Lu, H.S.; Martin, F.; Chang, M.
Cell 77, 1117-1124, 1994
A:Title: Identification and cloning of a megakaryocyte growth and development factor
A:Reference number: A54463; MUID:94291201; PMID:8020099
A:Accession: 138672
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-112, 'E', 114-353 <RE3>
A:Cross-references: EMBL:U1025; NID:g511223; PIDB:AAA50553.1; PID:g558078
R:Günery, A.L.; Xiang, W.J.; Xie, M.H.; Malloy, B.E.; Eaton, D.L.; de Sauvage, F.J.
Blood 85, 981-988, 1995
A:Title: Genomic structure, chromosomal localization, and conserved alternative splic
A:Reference number: I52610; MUID:95152076; PMID:7849339
A:Accession: 152610
A:Status: preliminary; translated from GR/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-353 <RE4>
A:Cross-references: GB:S76771; NID:g914225; PIDB:AAB33390.1; PID:g914226
A:Accession: S48740
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <SOH>
A:Cross-references: GB:D32046; NID:g5773319; PIDB:BAA06807.1; PID:g5773320
R:Batley, T.D.; Bogenberger, J.; Hunt, P.; Li, Y.S.; Lu, H.S.; Martin, F.; Chang, M.
Cell 77, 1117-1124, 1994
A:Title: Identification and cloning of a megakaryocyte growth and development factor
A:Reference number: A54463; MUID:94291201; PMID:8020099
A:Accession: 138672
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-112, 'E', 114-353 <RE3>
A:Cross-references: EMBL:U1025; NID:g511223; PIDB:AAA50553.1; PID:g558078
R:Günery, A.L.; Xiang, W.J.; Xie, M.H.; Malloy, B.E.; Eaton, D.L.; de Sauvage, F.J.
Blood 85, 981-988, 1995
A:Title: Genomic structure, chromosomal localization, and conserved alternative splic
A:Reference number: I52610; MUID:95152076; PMID:7849339
A:Accession: 152610
A:Status: preliminary; translated from GR/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-353 <RE4>
A:Cross-references: GB:S76771; NID:g914225; PIDB:AAB33390.1; PID:g914226
A:Accession: S48740
A:Status: preliminary
A:Molecule type: DNA

Query Match	9.7%	Score 96.5	DB 2	Length 353
Best Local Similarity	26.7%	Pred. No. 0.27		
Matches	46	Conservative	21	Mismatches 82; Indels 23; Gaps 6
QY	12	ILLSLSLPLGAPVIGAPRLICDSRVLERYLLLEAKAEANITTCGAHCSELENITVPDT	71	
DB	11	MULTIARLTLSSP--APP--ACDLRLYSKLLRDSRVLHLSRLSQCEVHPPLTTPVLLPRAV	65	
QY	72	KVNFAAKRMKEYRQQAQAVEYMGCLALLSEAVL--RCQALLVNSQPEWPELQLHVDKAVSCL	129	
DB	66	DESLSEEMTQMETEAKODILGAVTLLEGVNMAARQGLPCTCSLLGQSGGVRLTLGAL	125	
QY	130	RSLLTLRLALRQAKAEIASPPDAASAAPLRTITADIFRKLFRVYSNLRCKLK	181	
DB	126	QSL-----LGTQ-----LPQG-----RTTAHKDPNALFLSFQHLRLGRKVR	161	

```

RESULT 13
A:55530
megakaryocyte growth and development factor, long form - human
N:Alternate names: MPL ligand, long form
C:Species: Homo sapiens (man)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 07-May-1999
C:Accession: A55530
R:Chang, M.; McNlich, J.; Basu, R.; Shutter, J.; Hsu, R.; Perkins, C.; Mar, V.; Suggs
J. Biol. Chem. 270, 511-514, 1995
A:Title: Cloning and characterization of the human megakaryocyte growth and developme
A:Reference number: A55530; MUID:95122483; PMID:7822271
A:Accession: A55530
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-286 <CHA>
A:Cross-references: GB:U17071

```

C:Genetics:
A:Gene: MGDF
A:Map position: 3q26.3
C:Keywords: alternative splicing; cytokine

Query Match 9.4%; Score 93.5; DB 2; Length 286;
Best Local Similarity 27.1%; Pred. No. 0.39; Mismatches 82; Indels 23; Gaps 6;
Matches 46; Conservative 19; Mismatches 82; Indels 23; Gaps 6;

OY 12 LILSLISLPGLVLPGLPRLICDSRYLEAKEENITTCAGAEHCSLNEITVPPDT 71
DB 11 MLITLRLITLSSP---APP--ACDLRLVLSKLRSDSHVLSRSCQCEVHPPLTPVLLPAV 65
OY 72 KNEFYAMKREYVROQAVEVWQGLALISEAVL--RGQALLVNSSQPEWPIQLHYDKAVSGI 129
DB 66 DESLGEWKTQMEETKADIDILGAVTLLEGVMAARGOLGPTCLSSILQQLSGQVRLILGAL 125
OY 130 RSLITLRLRLGAKKEAISPDDAASAPLRITITADTRFKLFERYVSNFLRGLK 179
DB 126 QSL-----LGTQ-----LPPQG-----RTTAHKDPNAIFLSFQHLRLK 159

RESULT 14

S45330
thrombopoietin - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Dec-1994 #sequence_revision 17-Nov-1995 #text_change 08-Oct-1999
C:Accession: S45330
R:Lox, S.; Kaushtansky, K.; Holly, R.D.; Kuiper, J.L.; Lofton-Day, C.E.; Oort, P.J.; Gra-
unkard, D.; Ching, A.F.T.; Mathewes, S.L.; Bailey, M.C.; Forstrom, J.W.; Buddle, M.M.; C
Nature 369, 565-568, 1994
A:Title: Cloning and expression of murine thrombopoietin cDNA and stimulation of platelet
A:Reference number: S45330; MUID:94261207; PMID:8202158
A:Accession: S45330
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-356 <LOK>
A:Cross-references: GB:L34169; NID:9508540; PIDN:AAA40436.1; PID:9508541

Query Match 9.0%; Score 90; DB 2; Length 356;
Best Local Similarity 25.1%; Pred. No. 1.1; Mismatches 78; Indels 24; Gaps 6;

Matches 44; Conservative 29; Mismatches 78; Indels 24; Gaps 6;
OY 12 LILSLISLPGLVLPGLPRLICDSRYLEAKEENITTCAGAEHCSLNEITVPPDT 71
DB 6 LLLAAMLAVALRLTSSPYAPACDPRLNKLRLSDSHLSRSCQCPDVPDLSIPVLLPAV 65
OY 72 KNEFYAMKREYVROQAVEVWQGLALISEAVL--RGQALLVNSSQPEWPIQLHYDKAVSGI 129
DB 66 DESLGEWKTQMEETKADIDILGAVTLLEGVMAARGOLGPTCLSSILQQLSGQVRLILGAL 125
OY 130 RSLITLRLRLGAKKEAISPDDAASAPLRITITADTRFKLFERYVSNFLRGLK 181
DB 112 GOLSGOVRLILGALIGL-----GTQLPLOGRTTAHKDPNALFLSLQQLRGRKVR 161

RESULT 15

AE09539
Solute binding receptor protein [imported] - Salmonella enterica subsp. enterica serovar
C:Species: Salmonella enterica subsp. enterica serovar Typh
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001
C:Accession: AE09539
R:Parhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; PMID:11677608
A:Accession: AE09539
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-346 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD03169.1; PID:q16504804; GSPDB:GN00176
C:Genetics:
A:Gene: STY3952

Query Match 8.8%; Score 87.5; DB 2; Length 346;
Best Local Similarity 26.7%; Pred. No. 1.8; Mismatches 48; Indels 51; Gaps 9;
Matches 44; Conservative 22; Mismatches 48; Indels 51; Gaps 9;

OY 37 RYERLYLEAKEENITTCG--CAEHCSLNE--NITVPDTRKVNFAKREYVROQAVEVWQ 92
DB 217 RNLLQEMLEHHPDANVYASATIAEAAEGRNLTTPITIVSYL-----THQYTR 267
OY 93 GLALLSEAVLRGQALLVNSSQ--PWEPIQLHYDKAVSGLSLTLRLALGAQ--KEAISP 149
DB 268 GLK-----RGHILMALSDQAMQ-----GELATQSTIKVLQGPVPENISPP 309
OY 150 -----DAASAAPLRITITADTRFKLFERYVSNFLRGLKLYTGEA 187
DB 310 VLLITHNNADSARVRSLSPPGFRPY-----LYQYTSFA 344

Search completed: January 8, 2003, 05:14:55
Job time : 41.8802 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 00:55:32 ; Search time 13.4401 Seconds
(without alignments)
595.600 Million cell updates/sec

Title: US-09-813-775c-2

Perfect score: 998

Sequence: 1 MGVECPAWLWLLSLSLP.....NFLRGLKLYTGECRTGDR 193

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	991	99.3	193	1	EPO_HUMAN
2	899.5	90.1	192	1	EPO_MACRA
3	894.5	89.6	192	1	EPO_MACMU
4	793.5	79.5	192	1	EPO_FELCA
5	778.5	78.0	192	1	EPO_RAT
6	773.5	77.5	194	1	EPO_SHEEP
7	769.5	77.1	190	1	EPO_PIG
8	766.5	76.8	192	1	EPO_MOUSE
9	761.5	76.3	192	1	EPO_BOVIN
10	714.5	71.6	175	1	EPO_CANFA
11	124.5	12.5	352	1	TPO_CANFA
12	96.5	9.7	353	1	TPO_HUMAN
13	90	9.0	356	1	TPO_MOUSE
14	86.5	8.7	326	1	TPO_RAT
15	86.5	8.7	622	1	FAGG_HUMAN
16	85.5	8.6	897	1	EP15_MOUSE
17	82.5	8.3	3033	1	POLG_HCVJ8
18	82	8.2	1980	1	MY9B_RAT
19	79.5	8.0	907	1	GACS_PSESY
20	79	7.9	1007	1	SAL2_HUMAN
21	79	7.9	1089	1	IMB3_YEAST
22	78	7.8	896	1	CH60_BUCMP
23	77.5	7.7	896	1	EP15_HUMAN
24	77	7.7	551	1	CH60_BUCAP
25	77	7.7	778	1	RG12_MOUSE
26	75.5	7.6	381	1	MODD_MYCAV
27	75.5	7.6	3164	1	TEG0_HSV1
28	74.5	7.5	353	1	NADA_YERPE
29	74.5	7.5	388	1	TRA6_BURCE
30	74.5	7.5	543	1	CH60_BARBA
31	74.5	7.5	552	1	CH60_PSRST
32	74.5	7.5	830	1	GCL2_MOUSE
33	74.5	7.5	837	1	GCL2_HUMAN

34	74	7.4	552	1	CH60_COXBU	P19421 coxiella bu
35	73.5	7.4	468	1	T10A_HUMAN	O00220 homo sapien
36	73.5	7.4	3010	1	POLG_HCVJA	P26662 h genome po
37	73	7.3	342	1	TORT_ECO57	P58358 escherichia
38	73	7.3	342	1	TORT_ECOLI	P38683 escherichia
39	73	7.3	548	1	CH60_BUCAI	P25750 buchiera ap
40	73	7.3	3010	1	POLG_HCVJT	Q00269 h genome po
41	72.5	7.3	486	1	BAF1_KIOLA	P26375 kluyveromyc
42	72	7.2	224	1	MER1_HUMAN	O9um22 homo sapien
43	72	7.2	263	1	YH25_DEIRA	O83030 deinococcus
44	72	7.2	809	1	HRPB_ECOLI	P37024 escherichia
45	71.3	7.2	547	1	CH60_LEGPN	P26878 legionella

ALIGNMENTS

RESULT 1
EPO_HUMAN STANDARD: PRT: 193 AA.
AC P01588; O9UHA0; O9UEZ5; O9UDZ0;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Erythropoietin precursor (Epoetin).
GN EPO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85137899; PubMed=3838366;
RA Jacobs K., Shoemaker C., Rudersdorf R., Neill S.D., Kaufman R.J.,
RA Mutsaers A., Seehra J., Jones S.S., Hewick R., Fritsch E.F.,
RA Kawakita M., Shimizu T., Miyake T.,
RT "Isolation and characterization of genomic and cDNA clones of human
erythropoietin."
RL Nature 313:806-810(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86067948; PubMed=3865178;
RA Lin F.-K., Suggs S., Lin C.-H., Browne J.K., Smalling R., Egrie J.C.,
RA Chen K.K., Fox G.M., Martin F., Stabinsky Z., Badrawi S.M., Lai P.-H.,
RA Goldwasser E.,
RT "Cloning and expression of the human erythropoietin gene."
RL Proc. Natl. Acad. Sci. U.S.A. 82:7580-7584(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99018118; PubMed=9799793;
RA Gloeckner G., Scherer S., Schattevoy R., Boright A., Weber J.,
RA Tsui L.-C., Rosenthal A.,
RT "Large-scale sequencing of two regions in human chromosome 7q22:
analysis of 650 kb of genomic sequence around the EPO and CDT1 loci
reveals 17 genes."
RL Genome Res. 8:1060-1073(1998).
RN [4]
RP SEQUENCE FROM N.A.
RX Rupert J.L., Hochachka P.W.,
RT "Erythropoietin gene sequence in the Quechua, a high altitude native
population."
RN submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
RN [5]
RP SEQUENCE OF 58-193 FROM N.A., AND VARIANTS HEPATOCELLULAR CARCINOMA.
RX MEDLINE=9384593; PubMed=8396923;
RA Funakoshi A., Muta H., Bada T., Shimizu S.,
RT "Gene expression of mutant erythropoietin in hepatocellular
carcinoma."
RN Biochem. Biophys. Res. Commun. 195:717-722(1993).
RP [6]
RC SEQUENCE OF 28-193, AND DISULFIDE BONDS.
RX TISSUE=Urine;
MEDLINE=86140080; PubMed=3949763;

RA Lai P H, Everett R, Wang F F, Arakawa T, Goldwasser E;
RT "Structural characterization of human erythropoietin.";
RL J. Biol. Chem. 261:3116-3121(1986).
RN (7)
RP PRELIMINARY SEQUENCE OF 28-57.
RX MEDLINE=84135751; PubMed=6698989;
RA Yanagawa S, Hirade K, Ohnoka H, Sasaki R, Chiba H, Ueda M,
RA Goto M;
RT "Isolation of human erythropoietin with monoclonal antibodies.";
RL J. Biol. Chem. 259:2707-2710(1984).
RN (8)
RP STRUCTURE OF CARBOHYDRATES.
RX MEDLINE=88153657; PubMed=3346214;
RA Takeuchi M, Takasaki S, Miyazaki H, Kato T, Hoshi S, Kochibe N,
RA Kobata A;
RT "Comparative study of the asparagine-linked sugar chains of human
RT erythropoietins purified from urine and the culture medium of
RT recombinant Chinese hamster ovary cells.";
RL J. Biol. Chem. 263:3657-3663(1988).
RN (9)
RP STRUCTURE OF CARBOHYDRATES.
RX MEDLINE=89118279; PubMed=3219367;
RA Sasaki H, Ochi N, Dell A, Fukuda M;
RT "Site-specific glycosylation of human recombinant erythropoietin:
RT analysis of glycopeptides or peptides at each glycosylation site by
RT fast atom bombardment mass spectrometry.";
RL Biochemistry 27:8618-8626(1988).
RN (10)
RP STRUCTURE OF CARBOHYDRATES.
RX MEDLINE=92314463; PubMed=1820196;
RA Takeuchi M, Kobata A;
RT "Structures and functional roles of the sugar chains of human
RT erythropoietins.";
RL Glycobiology 1:337-346(1991).
RN (11)
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=98445092; PubMed=9774108;
RA Syed R S, Reid S W, Li C, Cheetham J C, Aoki K H, Liu B,
RA Zhan H, Ostlund T D, Chittino A J, Zhang J, Finner-Moore J,
RA Elliott S, Stiney K, Katz B A, Matthews D J, Wendoloski J J,
RA Egrie J, Stroud R M;
RT "Efficiency of signalling through cytokine receptors depends
RT critically on receptor orientation.";
RL Nature 395:511-516(1998).
CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC -1- PHARMACOLOGICAL: Used for the treatment of anemia. Available under
CC the names EPOgen (Amgen), EPOgin (Chugai), Epomax (Eli Lilly), Eprex
CC (Janssen-Cilag), Neorecormon or Recormon (Roche), and Procrit
CC (Ortho Biotech). Variations in the glycosylation pattern of EPO
CC distinguishes these products. EPOgen, EPOgin, Eprex and Procrit
CC are generically known as epoetin alfa, Neorecormon and Recormon as
CC epoetin beta and Epomax as epoetin omega.
CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
CC -1- DATABASE: NAME=Rad Systems' cytokine source book: EPO;
CC WWW="http://www.rndsystems.com/aap/g_sitebuilder.asp?bodyId=197".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as the content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X02158; CAA26095.1; -
DR EMBL: X02157; CAA26094.1; -
DR EMBL: M1319; AAA52400.1; -
DR EMBL: AF053356; AAC78791.1; -

DR	EMBL: AF020308; AAF23132.1; -	DR	EMBL: AF020306; AAF23132.1; JOINED.
DR	EMBL: AF020307; AAF23132.1; JOINED.	DR	EMBL: AF020307; AAF23132.1; JOINED.
DR	EMBL: AF020310; AAF23133.1; -	DR	EMBL: AF020310; AAF23133.1; -
DR	EMBL: AF020309; AAF23133.1; JOINED.	DR	EMBL: AF020309; AAF23133.1; JOINED.
DR	EMBL: AF020311; AAF15752.1; -	DR	EMBL: AF020311; AAF15752.1; -
DR	EMBL: AF020314; AAF23134.1; -	DR	EMBL: AF020314; AAF23134.1; -
DR	EMBL: AF020312; AAF23134.1; JOINED.	DR	EMBL: AF020312; AAF23134.1; JOINED.
DR	EMBL: AF020313; AAF23134.1; JOINED.	DR	EMBL: AF020313; AAF23134.1; JOINED.
DR	EMBL: S65458; AAD13964.1; -	DR	EMBL: S65458; AAD13964.1; -
DR	PIR: A01855; ZUHU.	DR	PIR: A01855; ZUHU.
DR	PIR: A25384; A25384.	DR	PIR: A25384; A25384.
DR	PIR: A24744; A24744.	DR	PIR: A24744; A24744.
DR	PIR: A22210; A22210.	DR	PIR: A22210; A22210.
DR	PDB: 1C94; 1I-AUG-99.	DR	PDB: 1C94; 1I-AUG-99.
DR	GlycosultedB: P01588; -	DR	GlycosultedB: P01588; -
DR	GeneW: HGNC:3415; EPO.	DR	GeneW: HGNC:3415; EPO.
DR	MIM: 133170; -	DR	MIM: 133170; -
DR	InterPro: IPR001323; EPO_TPO.	DR	InterPro: IPR001323; EPO_TPO.
DR	InterPro: IPR003013; Erythroptn.	DR	InterPro: IPR003013; Erythroptn.
DR	Pfam: PF00758; EPO_TPO; 1.	DR	Pfam: PF00758; EPO_TPO; 1.
DR	PRINTS: PR00272; ERYTHROPTN.	DR	PRINTS: PR00272; ERYTHROPTN.
DR	PROSITE: PS00817; EPO_TPO; 1.	DR	PROSITE: PS00817; EPO_TPO; 1.
KW	Erythrocyte maturation; Glycoprotein; Hormone; Signal; Pharmaceutical; 3D-structure.	KW	Erythrocyte maturation; Glycoprotein; Hormone; Signal; Pharmaceutical; 3D-structure.
FT	SIGNAL 1 27	FT	SIGNAL 1 27
FT	CHAIN 28 193	FT	CHAIN 28 193
FT	PROPE 190 193	FT	PROPE 190 193
FT	DISULFID 34 188	FT	DISULFID 34 188
FT	DISULFID 56 60	FT	DISULFID 56 60
FT	CARBOHYD 51 51	FT	CARBOHYD 51 51
FT	CARBOHYD 65 65	FT	CARBOHYD 65 65
FT	CARBOHYD 110 110	FT	CARBOHYD 110 110
FT	CARBOHYD 153 153	FT	CARBOHYD 153 153
FT	VARIANT 131 132	FT	VARIANT 131 132
FT	VARIANT 149 149	FT	VARIANT 149 149
FT	CONFLICT 40 40	FT	CONFLICT 40 40
FT	CONFLICT 85 85	FT	CONFLICT 85 85
FT	CONFLICT 140 140	FT	CONFLICT 140 140
SQ	SEQUENCE 193 AA; 21306 MW; C91F0E4C26A52033 CRC64;	SQ	SEQUENCE 193 AA; 21306 MW; C91F0E4C26A52033 CRC64;

Query Match 99.3%; Score 991; DB 1; Length 193;
Best Local Similarity 99.5%; Pred. No. 4, 4e-87;
Matches 192; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY	1	MGVHECPAMWMLLSLPLGLPIVGPAPRLICSRVLEKLEAKENITGCAEHC	60
DB	1	MGVHECPAMWMLLSLPLGLPIVGPAPRLICSRVLEKLEAKENITGCAEHC	60
OY	61	SLNENITVDPTKVFNFYAMKRMVEVROQAVEVMGGLLSAVLRGQALLVNSSQPEPDLQ	120
DB	61	SLNENITVDPTKVFNFYAMKRMVEVROQAVEVMGGLLSAVLRGQALLVNSSQPEPDLQ	120
OY	121	HYDKAVSGRLSTLLRALGAOKREKISPDAASAPLRTITDTERKLEFRVYSNFIKGL	180
DB	121	HYDKAVSGRLSTLLRALGAOKREKISPDAASAPLRTITDTERKLEFRVYSNFIKGL	180
OY	181	KLYTGEACRTGDR	193
DB	181	KLYTGEACRTGDR	193

RESULT 2
EPO_MACFA
ID EPO_MACFA STANDARD; PRT: 192 AA.
AC P07865;

DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, last sequence update)
 DT 16-OCT-2001 (Rel. 40, last annotation update)
 DE Erythropoietin precursor.
 GN EPO.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 NC NCBI_Taxid=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=67055236; PubMed=2877922;
 RA Lin F.-K., Lin C.-H., Lai P.-H., Browne J.K., Egrie J.C., Smalling R.,
 RA Fox G.M., Chen K.K., Castro M., Suggs S.;
 RT "Monkey erythropoietin gene: cloning, expression and comparison with
 RT the human erythropoietin gene.";
 RL Gene 44:201-209(1986)
 CC -!- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
 CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
 CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
 CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
 CC -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M18189; AAA36841.1; -;
 DR PIR: J00173; J00173.
 DR HSSP: P01588; ICN4.
 DR InterPro: IPR001323; EPO_TPO.
 DR Pfam: PF00758; EPO_TPO.1.
 DR PRINTS: PR00272; ERYTHROPTN.
 DR PROSITE: PS00817; EPO_TPO.1.
 DR Erythrocyte maturation; Glycoprotein; Hormone; Signal.
 KM SIGNAL 1 27
 FT CHAIN 28 192
 FT DISULFID 34 187
 FT FT 56 60
 FT CARBOHYD 51 51
 FT CARBOHYD 65 65
 FT CARBOHYD 110 110
 FT CARBOHYD 152 152
 SQ SEQUENCE 192 AA; 21113 MW; E8A900F442AD4522 CRC64;
 Query Match. 90.1%; Score 899.5; DB 1; Length 192;
 Best Local Similarity 91.2%; Pred. No. 2.2e-78;
 Matches 176; Conservative 8; Mismatches 8; Indels 1; Gaps 1;
 QY 1 MGVECPAMWMLLSLPLGLPVLGAPAPRLICDSRYLERYLLEAKAEENITGCAEHC 60
 DB 1 MGVECPAMWMLLSLPLGLPVLGAPAPRLICDSRYLERYLLEAKAEENITGCAEHC 60
 DB 61 SLNENITVPDTKVNFAWKREVRQOAVEVWQGLALLSEAVLRQALLVNSQFPEPLQ 120
 QY 121 HVDKAVSGLRSLTTLRLALGOKAISPDAASAAPLRTTAQDFRRLFRYSNFKGL 180
 DB 121 HMDKAISGLRSITTLRLALGQO-EAISLPDAASAAPLRTTAQDFRRLFRYSNFKGL 179
 QY 181 KLYTGACRTGDR 193
 DB 180 KLYTGACRTGDR 192

RESULT 3
 EPO_MACMU
 ID EPO_MACMU STANDARD; PRT; 192 AA.
 AC 028513;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, last sequence update)
 DT 16-OCT-2001 (Rel. 40, last annotation update)
 DE Erythropoietin precursor.
 GN EPO.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 NC NCBI_Taxid=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93372347; PubMed=8364201;
 RA Wen D., Boissel J.P.R., Tracy T.E., Gruninger R.H., Mulcahy L.S.,
 RA Czelusniak J., Goodman M., Bunn H.F.;
 RT "Erythropoietin structure-function relationships: high degree of
 RT sequence homology among mammals.";
 RL Blood 82:1507-1516(1993).
 CC -!- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
 CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
 CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
 CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
 CC -!- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L10609; AAA36842.1; -;
 DR HSSP: P01588; ICN4.
 DR InterPro: IPR001323; EPO_TPO.
 DR Pfam: PF00758; EPO_TPO.1.
 DR PRINTS: PR00272; ERYTHROPTN.
 DR PROSITE: PS00817; EPO_TPO.1.
 DR Erythrocyte maturation; Glycoprotein; Hormone; Signal.
 KM SIGNAL 1 27
 FT CHAIN 28 192
 FT DISULFID 34 187
 FT FT 56 60
 FT CARBOHYD 51 51
 FT CARBOHYD 65 65
 FT CARBOHYD 110 110
 FT CARBOHYD 152 152
 SQ SEQUENCE 192 AA; 21081 MW; 275560A264628CD1 CRC64;
 Query Match. 89.6%; Score 894.5; DB 1; Length 192;
 Best Local Similarity 90.2%; Pred. No. 6.4e-78;
 Matches 174; Conservative 10; Mismatches 8; Indels 1; Gaps 1;
 QY 1 MGVECPAMWMLLSLPLGLPVLGAPAPRLICDSRYLERYLLEAKAEENITGCAEHC 60
 DB 1 MGVECPAMWMLLSLPLGLPVLGAPAPRLICDSRYLERYLLEAKAEENITGCAEHC 60
 DB 61 SLNENITVPDTKVNFAWKREVRQOAVEVWQGLALLSEAVLRQALLVNSQFPEPLQ 120
 QY 121 HVDKAVSGLRSLTTLRLALGOKAISPDAASAAPLRTTAQDFRRLFRYSNFKGL 180
 DB 121 HMDKAISGLRSITTLRLALGQO-EAISLPDAASAAPLRTTAQDFRRLFRYSNFKGL 179

```

Db      1 MGSECCP-LILLLSLILLPLGLPLGAPPRCLICDSRLVERYLICAREAKENTMCAECG 59
QY      61 SLNENITVPTDKRVNFAKAKREVRQQAEEVWOGALLSEAVIRGALLVNSQPHEPQL 120
Db      60 SFSEKITVPTDKRVNFTYTKKRDVGGAAVEWOGALLSEAILRGALLANSQPSETIQL 119
QY      121 HDKAVSGLRSITLTLRALGAQKEAISPDDASAAPLRTITADPFRKLFVRSNRLRGL 180
Db      120 HDKAVSSLRSLTSLRALGAQKEATSLPEATSAAPLRTFTYDICTKLFRTSNRLRGL 179
QY      181 KLYTGEACRTGDR 193
Db      180 TLYTGEACRGRDR 192

RESULT 5
EPO_RAT
ID      EPO_RAT      STANDARD;      PRT;      192 AA.
AC      P29676; P70504;
DT      01-APR-1993 (Rel. 25, Created)
DT      01-APR-1993 (Rel. 25, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Erythropoietin precursor.
GN      EPO.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxId=10116;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Mistat; TISSUE=Kidney;
RX      MEDLINE=93042015; PubMed=1420369;
RA      Nagao M., Suga H., Okano M., Masuda S., Narita H., Ikura K.,
RA      Sasaki R.;
RT      "Nucleotide sequence of rat erythropoietin.";
RL      Biochim. Biophys. Acta 1171:99-102(1992).
[2]
RP      SEQUENCE OF 4-192 FROM N.A.
RC      STRAIN=Sprague-Dawley; TISSUE=Kidney;
RX      MEDLINE=93372347; PubMed=8364201;
RA      Wen D., Bolsens J.P.R., Tracy T.E., Mulcahy L.S., Czelusniak J.,
RA      Goodman M., Bunn H.F.;
RT      "Erythropoietin structure-function relationships: high degree of
RT      sequence homology among mammals.";
RL      Blood 82:1507-1516(1993).
-1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
-1- REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
-1- PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
-1- SUBCELLULAR LOCATION: Secreted.
-1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
-1- AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
-1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
-----
CC      CC
CC      CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      CC the European Bioinformatics Institute. There are no restrictions on its
CC      CC use by non-profit institutions as long as its content is in no way
CC      CC modified and this statement is not removed. Usage by and for commercial
CC      CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      CC or send an email to license@isb-sib.ch).
-----
CC      CC
CC      CC EMBL; D10763; BAA01593.1; -
CC      CC EMBL; L10608; AAA1126.1; -
CC      CC PIR; S28148; S28148.
CC      CC DR
CC      CC HSP; P01588; ICN4.
CC      CC DR
CC      CC InterPro; IPR001323; EPO_TPO.
CC      CC DR
CC      CC Pfam; PF00758; EPO_TPO; 1.
CC      CC DR
CC      CC PRINTS; PR00272; ERYTHROPEN.
CC      CC DR
CC      CC PROSITE; PS00817; EPO_TPO; 1.
CC      CC KM
CC      CC Erythrocyte maturation; Glycoprotein; Hormone; Signal.
CC      CC FT
CC      CC SIGNAL 1 26 BY SIMILARITY.
CC      CC CHAIN 27 192 ERYTHROPOIETIN.

```


This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: L10607; AAA31029.1; -
 CC HSSP: P01588; 1CN4.
 DR InterPro: IPR001323; EPO_TPO.
 DR Pfam: PF00758; EPO_TPO; 1.
 DR PROSITE: PS00817; EPO_TPO; 1.
 KM Erythrocyte maturation; Glycoprotein; Hormone; Signal.
 FT MON_TER 1 1
 FT SIGNAL <1 22 POTENTIAL.
 FT CHAIN 23 190 ERYTHROPOIETIN.
 FT DISULFID 29 185 BY SIMILARITY.
 FT DISULFID 51 55
 FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 190 AA; 20888 MW; A75BD6CCE5077E2A CRC64;

Query Match 77.1%; Score 769.5; DB 1; Length 190;
 Best Local Similarity 82.7%; Pred. No. 4,8e-66;
 Matches 158; Conservative 8; Mismatches 22; Indels 3; Gaps 2;

OY 5 ECPAWMLLTLSTLPLGLPVLCAPRLICDSRVLEERYLEEAKENITGCAEHGSLN 64
 DB 1 ECPARL-LTLSTLPLGLPVLCAPRLICDSRVLEERYLEEAKENITGCAEHGSLN 59
 OY 65 NITVPDTKYVNFYAMKREYVQQAWEVWGLALSLSEAVLRGQALLVNSQPWEPLQLHVDK 124
 DB 60 NITVPDTKYVNFYAMKREYVQQAWEVWGLALSLSEAVLRGQALLVNSQPWEPLQLHVDK 119
 OY 125 AVSGRLSTLTLRALGAKKEAISPPDA--ASAAPLRTTADTFKRLFRYVSNFLRKRLK 182
 DB 120 AVSGRLSTLTLRALGAKKEAISPPDA--ASAAPLRTTADTFKRLFRYVSNFLRKRLK 179
 OY 183 YTMGEACRTGDR 193
 DB 180 YTMGEACRTGDR 190
 RESULT 8
 EPO_MOUSE
 ID EPO_MOUSE STANDARD; PRT; 192 AA.
 AC P07321;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Erythropoietin precursor.
 GN EPO.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP MEDLINE=87039105; PubMed=3773894;
 RA Shoemaker C.B., Mltsock L.D.;
 RT "Murine erythropoietin gene: cloning, expression, and human gene
 RT homology.";
 RL Mol. Cell. Biol. 6:849-858(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=87039104; PubMed=3022133;
 RA McDonald J.D., Lin F.-K., Goldwasser E.;
 RT "Cloning, sequencing, and evolutionary analysis of the mouse
 RT erythropoietin gene.";
 RL Mol. Cell. Biol. 6:842-848(1986).

RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SV;
 RX MEDLINE=21138439; PubMed=11239002;
 RA Wilson M.D., Riemer C., Martindale D.W., Schnupf P., Boright A.P.,
 RA Cheung T.L., Hardy D.M., Schwartz S., Scherer S.W., Tsui L.-C.,
 RA Miller W., Koop B.F.;
 RT "Comparative analysis of the gene-dense Ache/TFE2 region on human
 RT chromosome 7q22 with the orthologous region on mouse chromosome 5.";
 RL Nucleic Acids Res. 29:1352-1365(2001).
 RN [4]

RP SEQUENCE OF 1-52 FROM N.A.
 RC STRAIN=ICW;
 RX MEDLINE=98030528; PubMed=9365246;
 RA Chretien S., Duprez V., Maouche L., Gisselbrecht S., Mayeux P.,
 RA Lacombe C.;
 RT "Abnormal erythropoietin (Epo) gene expression in the murine
 RT erythroleukemia Iw32 cells results from a rearrangement between the
 RT G-protein beta2 subunit gene and the Epo gene.";
 RL Oncogene 15:1995-1999(1997).
 CC -1- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
 CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
 CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
 CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
 CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M12482; AAA37568.1; -
 DR EMBL: M12950; AAA37570.1; -
 DR EMBL: AF312033; AAK28825.1; -
 DR EMBL: Y11971; CAA272707.1; -
 DR PIR: A24901; A24901.
 DR PIR: A24902; A24902.
 DR HSSP: P01588; 1CN4.
 DR MGD: MGI:95407; EPO.
 DR InterPro: IPR001323; EPO_TPO.
 DR InterPro: IPR003013; Erythroptn.
 DR Pfam: PF00758; EPO_TPO; 1.
 DR PRINTS: PR00272; ERYTHROPTN.
 DR PROSITE: PS00817; EPO_TPO; 1.
 KM Erythrocyte maturation; Glycoprotein; Hormone; Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 192 ERYTHROPOIETIN.
 FT DISULFID 33 187 BY SIMILARITY.
 FT CARBOHYD 50 50 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
 FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
 FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
 SQ SEQUENCE 192 AA; 21365 MW; 65F94E214E0DEF2E CRC64;

Query Match 76.8%; Score 766.5; DB 1; Length 192;
 Best Local Similarity 78.8%; Pred. No. 9.4e-66;
 Matches 152; Conservative 15; Mismatches 25; Indels 1; Gaps 1;

OY 1 MGVHECPAWMLLTLSTLPLGLPVLCAPRLICDSRVLEERYLEEAKENITGCAEHG 60
 DB 1 MGVHECPAWMLLTLSTLPLGLPVLCAPRLICDSRVLEERYLEEAKENITGCAEHG 59
 OY 61 SLNENITVPDTKYVNFYAMKREYVQQAWEVWGLALSLSEAVLRGQALLVNSQPWEPLQL 120
 DB 60 SLNENITVPDTKYVNFYAMKREYVQQAWEVWGLALSLSEAVLRGQALLVNSQPWEPLQL 119
 OY 121 HYDKAVSGRLSTLTLRALGAKKEAISPPDA--ASAAPLRTTADTFKRLFRYVSNFLRKRL 180
 DB 120 HYDKAVSGRLSTLTLRALGAKKEAISPPDA--ASAAPLRTTADTFKRLFRYVSNFLRKRL 179

OY 65 NITVPDTKVFYAKRMEVQOAVEVQOGLALLSEVLRQALLVNSQWPELOLHVDR 124
 DB 60 NITVPDTKVFYAKRMEVQOAVEVQOGLALLSEVLRQALLVNSQWPELOLHVDR 119
 OY 125 AVSGLSLTLLRALGAKQKAEISPPDASAPLRTTADTFKRLFRVNSFLR 180
 DB 120 AVSGLSLTLLRALGAKQKAEISPPDASAPLRTTADTFKRLFRVNSFLR 175

RESULT 11
 TPO_CANFA
 ID TPO_CANFA STANDARD: PRT: 352 AA.

AC P42705;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Thrombopoietin precursor (Megakaryocyte colony stimulating factor)
 DE (C-MPL ligand) (MPL) (Megakaryocyte growth and development factor)
 DE (MGDF).
 GN THPO OR TPO.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 24-44.
 RC TISSUE=Kidney;
 RA Bartley T.D., Bogenberger J., Hunt P., Li Y.-S., Lu H.-S., Martin F.,
 Chang M.-S., Samal B.B., Nichol J.L., Swift S., Johnson M.J.,
 Hsu R.-Y., Parker V.P., Suggs S., Skrine J.D., Merewether L.A.,
 Clogson C., Hsu E., Hokom M.M., Hornkohl A., Choi E., Pangelinan M.,
 Sun Y., Mar V., McNich J., Simonet L., Jacobsen F., Xie C.,
 Shutter J., Chute H., Basu R., Selander L., Trollinger D., Siew L.,
 Padilla D., Trail G., Elliott G., Izumi R., Covey T., Crouse J.,
 Garcia A., Xu W., del Castillo J., Biron J., Cole S., Hu M.C.-T.,
 Pacifici R., Ponting I., Sarris C., Wen D., Yung Y.P., Lin H.,
 Rosseman R.A.;
 RA "Identification and cloning of a megakaryocyte growth and development
 factor that is a ligand for the cytokine receptor Mpl.";
 RT Cell 77:1117-1117(1994).
 RL [1]
 CC -1- FUNCTION: LINEAGE-SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION
 AND MATURATION OF MEGAKARYOCYTES FROM THEIR COMMITTED PROGENITOR
 CELLS. IT ACTS AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT
 MAY BE THE MAJOR PHYSIOLOGICAL REGULATOR OF CIRCULATING PLATELETS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DOMAIN: TWO-DOMAIN STRUCTURE WITH AN ERYTHROPOIETIN-LIKE N-
 TERMINAL AND A SER/PRO/THR-RICH C-TERMINAL.
 CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
 DR InterPro: IPR001323; EPO_TPO.
 DR InterPro: IPR003978; Thrombopoietin.
 DR Pfam: PF00758; EPO_TPO; 1.
 DR PRINTS: PRO1485; THROMBOPTN.
 DR PROSITE: PS00817; EPO_TPO; 1.
 KW Cytokine; Glycoprotein; Hormone; Signal.

FT CHAIN 1 23
 FT SIGNAL 352
 FT CHAIN 24 352
 FT DISULFID 28 172
 FT DISULFID 50 106
 FT CARBOHYD 185 185
 FT CARBOHYD 197 197
 FT CARBOHYD 206 206
 FT CARBOHYD 234 234
 FT CARBOHYD 255 255
 FT CARBOHYD 332 332
 FT CARBOHYD 347 347
 SO SEQUENCE 352 AA; 37641 MW; 024F3B41B061PBD8 CR664;

Query Match 12.5%; Score 124.5; DB 1; Length 352;
 Best Local Similarity 25.5%; Pred. No. 0.0002;
 Matches 47; Conservative 26; Mismatches 68; Indels 43; Gaps 6;

OY 12 LLSLSLPLVLGL-PVLGAPRLICDSRVLEKRLLEAKKAEINTTGCACRHCNSMENTTVPD 70

DB 7 LLSVMTLLTARLDPCCPAP--ACDPRLLNKMRLDSVLSRLSQCDDIYPLSTVLLPA 64
 OY 71 TKVNFYAKRMEVQOAVEVQOGLALLSEVLRQALLVNSQWPELOLHVDR 128
 DB 65 VDSLSGKWKTKQKQTRAKQDVGAVALLDGLVLRGOL-----G 103
 OY 129 LRSITLTLRALGAKQKAEI-----SPPDASAPLRTTADTFKRLFRVNSFLR 177
 DB 104 PCLSSLLGLSGQVRLLAGLGLGLTQLPQG-----RTTHDPNMFILSPQOLLR 157
 OY 178 GKLR 181
 DB 158 GKVR 161

RESULT 12
 TPO_HUMAN
 ID TPO_HUMAN STANDARD: PRT: 353 AA.
 AC P40225; Q13020; Q15790; Q15791; Q15792;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Thrombopoietin precursor (Megakaryocyte colony stimulating factor)
 DE (Myeloproliferative leukemia virus oncogene ligand) (C-mpl ligand)
 DE (ML) (Megakaryocyte growth and development factor) (MGDF).
 GN THPO.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Fetal liver;
 RA MEDLINE=94261202; PubMed=8020154;
 RA de Sauvage F.J., Hass P.E., Spencer S.D., Malloy B.E., Gurney A.L.,
 Oles K.J., Hultgren B., Solberg L.A., Jr., Goeddel D.V., Eaton D.L.;
 RA "Stimulation of megakaryocytopoiesis and thrombopoiesis by the c-Mpl
 ligand.";
 RT Nature 369:533-538(1994).
 RL [2]
 CC -1- FUNCTION: LINEAGE-SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION
 AND MATURATION OF MEGAKARYOCYTES FROM THEIR COMMITTED PROGENITOR
 CELLS. IT ACTS AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT
 MAY BE THE MAJOR PHYSIOLOGICAL REGULATOR OF CIRCULATING PLATELETS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DOMAIN: TWO-DOMAIN STRUCTURE WITH AN ERYTHROPOIETIN-LIKE N-
 TERMINAL AND A SER/PRO/THR-RICH C-TERMINAL.
 CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
 DR InterPro: IPR001323; EPO_TPO.
 DR InterPro: IPR003978; Thrombopoietin.
 DR Pfam: PF00758; EPO_TPO; 1.
 DR PRINTS: PRO1485; THROMBOPTN.
 DR PROSITE: PS00817; EPO_TPO; 1.
 KW Cytokine; Glycoprotein; Hormone; Signal.

FT CHAIN 1 23
 FT SIGNAL 353
 FT CHAIN 24 353
 FT DISULFID 28 172
 FT DISULFID 50 106
 FT CARBOHYD 185 185
 FT CARBOHYD 197 197
 FT CARBOHYD 206 206
 FT CARBOHYD 234 234
 FT CARBOHYD 255 255
 FT CARBOHYD 332 332
 FT CARBOHYD 347 347
 SO SEQUENCE FROM N.A. (ISOFORM 1).
 MEDLINE=95108091; PubMed=7809166;
 RA Foster D.C., Sprecher C.A., Grant F.J., Kramer J.M., Kulper J.L.,
 Holly R.D., Whitmore T.E., Heibel M.D., Bell L.A.N., Ching A.F.,
 McGee V., Hart C., O'Hara P.J., Lok S.;
 RA "Human thrombopoietin: gene structure, cDNA sequence, expression, and
 chromosomal localization.";
 RT Proc. Natl. Acad. Sci. U.S.A. 91:13023-13027(1994).
 RL [4]
 CC -1- FUNCTION: LINEAGE-SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION
 AND MATURATION OF MEGAKARYOCYTES FROM THEIR COMMITTED PROGENITOR
 CELLS. IT ACTS AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT
 MAY BE THE MAJOR PHYSIOLOGICAL REGULATOR OF CIRCULATING PLATELETS.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DOMAIN: TWO-DOMAIN STRUCTURE WITH AN ERYTHROPOIETIN-LIKE N-
 TERMINAL AND A SER/PRO/THR-RICH C-TERMINAL.
 CC -1- SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
 DR InterPro: IPR001323; EPO_TPO.
 DR InterPro: IPR003978; Thrombopoietin.
 DR Pfam: PF00758; EPO_TPO; 1.
 DR PRINTS: PRO1485; THROMBOPTN.
 DR PROSITE: PS00817; EPO_TPO; 1.
 KW Cytokine; Glycoprotein; Hormone; Signal.

Query Match 12.5%; Score 124.5; DB 1; Length 352;
 Best Local Similarity 25.5%; Pred. No. 0.0002;
 Matches 47; Conservative 26; Mismatches 68; Indels 43; Gaps 6;
 OY 12 LLSLSLPLVLGL-PVLGAPRLICDSRVLEKRLLEAKKAEINTTGCACRHCNSMENTTVPD 70

RC SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=LIVER;
 RA MEDLINE=94291201; PubMed=8020099;
 RA Bartley T.D., Bogenberger J., Hunt P., Li Y.-S., Lu H.-S., Martin F.,
 RA Chang M.-S., Samal B.B., Nichol J.L., Swift S., Johnson M.J.,
 RA Hsu R.-Y., Parker V.P., Suggs S., Skrine J.D., Merewether L.A.,
 RA Clogson C., Hsu E., Hokom M.M., Hornkohl A., Choi E., Pangelinan M.,
 RA Sun Y., Mar V., Kenich J., Simonet L., Jacobsen F., Xie C.,
 RA Shutter J., Chute H., Basu R., Selander L., Trollinger D., Siew L.,
 RA Padilla D., Trail G., Elliott G., Izumi R., Covey T., Crose J.,
 RA Garcia A., Xu W., del Castillo J., Biron J., Cole S., Hu M.C.-T.,
 RA Pacifici R., Ponting I., Sarris C., Wen D., Yung Y.P., Lin H.,
 RA Bosseman R.A.;
 RT Identification and cloning of a megakaryocyte growth and development
 RT factor that is a ligand for the cytokine receptor MPL.;
 RT Cell 77:1117-1117(1994).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RP MEDLINE=95152076; PubMed=7849319;
 RX Gurney A.L., Kuang W.-J., Xie M.-H., Malloy B.E., Eaton D.L.,
 RA de Sauvage F.J.;
 RA Genomic structure, chromosomal localization, and conserved
 RT alternative splice forms of thrombopoietin.;
 RT Blood 85:981-988(1995).
 CC - FUNCTION: LINEAGE-SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION
 CC AND MATURATION OF MEGAKARYOCYTES FROM THEIR COMMITTED PROGENITOR
 CC CELLS. IT ACTS AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT
 CC MAY BE THE MAJOR PHYSIOLOGICAL REGULATOR OF CIRCULATING PLATELETS.
 CC - SUBCELLULAR LOCATION: Secreted.
 CC - ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1 (SHOWN HERE) AND 2/TPO-2, ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC - TISSUE SPECIFICITY: FOUND MAINLY IN THE LIVER, KIDNEY AND SKELETAL
 CC MUSCLE.
 CC - DOMAIN: TWO-DOMAIN STRUCTURE WITH AN ERYTHROPOIETIN-LIKE N-
 CC TERMINAL AND A SER/PRO/THR-RICH C-TERMINAL.
 CC - SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U34169; AAA0436.1; -;
 DR PIR: S45330; S45330.
 DR MGD: MGI:101875; Thpo.
 DR InterPro: IPR001323; EPO_TPO.
 DR InterPro: IPR003978; thrombopoietin.
 DR Pfam: PF00758; EPO_TPO; 1.
 DR PRINTS: PR01485; THROMBOPTN.
 DR PROSITE: PS00817; EPO_TPO; 1.
 KW Cytokine; Glycoprotein; Hormone; Signal; Alternative splicing.
 FT SIGNAL 1 21
 FT CHAIN 1 22
 FT DISULFID 28 356
 FT DISULFID 50 172
 FT CARBOHYD 197 197
 FT CARBOHYD 206 206
 FT CARBOHYD 235 235
 FT CARBOHYD 249 249
 FT CARBOHYD 256 256
 FT CARBOHYD 266 266
 FT CARBOHYD 336 336
 FT CARBOHYD 351 351
 FT VARSPIC 133 136
 SO SEQUENCE 356 AA; 37835 MW; D346DBAE8392053C CRC64;

DB 6 L1LAAMLAARLTLSPPAPACDPRLNKLNDLSLSRLSQCPDVPDLSPVLLPAV 65
 OY 72 KANFYAKRMREYQOAVVEWOGUALLSEAVL--RGQALVNSSQPMPEQLJHDKAVSG 129
 DB 66 DFLGCKKQYOTBQSKADILGAVSLLEGVMAKRGQ-----EPSCLS-----SL 111
 OY 130 RSITTLRL-ALGAQKEAISPDAASAPL--RTTADFEKLEPRVSNFLRGK 181
 DB 112 GOLSGVRLILGALGILL-----GTOLPLOGRTTAKRDPALFLSLQQLRGKVR 161
 RESULT 14
 ID TPO_RAT STANDARD; PRT; 326 AA.
 AC P49745;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Thrombopoietin precursor.
 GN Thpo.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE=95331639; PubMed=7607561;
 RA Ogami K., Shimada Y., Sohma Y., Akahori H., Kato T., Kawamura K.,
 RA Miyazaki H.;
 RT The sequence of a rat cDNA encoding thrombopoietin.;
 RL Gene 158:309-310(1995).
 CC - FUNCTION: LINEAGE-SPECIFIC CYTOKINE AFFECTING THE PROLIFERATION
 CC AND MATURATION OF MEGAKARYOCYTES FROM THEIR COMMITTED PROGENITOR
 CC CELLS. IT ACTS AT A LATE STAGE OF MEGAKARYOCYTE DEVELOPMENT. IT
 CC MAY BE THE MAJOR PHYSIOLOGICAL REGULATOR OF CIRCULATING PLATELETS.
 CC - SUBCELLULAR LOCATION: Secreted.
 CC - DOMAIN: TWO-DOMAIN STRUCTURE WITH AN ERYTHROPOIETIN-LIKE N-
 CC TERMINAL AND A SER/PRO/THR-RICH C-TERMINAL.
 CC - SIMILARITY: BELONGS TO THE EPO / TPO FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D32207; BAA06906.1; -;
 DR InterPro: IPR001323; EPO_TPO.
 DR InterPro: IPR003978; thrombopoietin.
 DR Pfam: PF00758; EPO_TPO; 1.
 DR PRINTS: PR01485; THROMBOPTN.
 DR PROSITE: PS00817; EPO_TPO; 1.
 KW Cytokine; Glycoprotein; Hormone; Signal.
 FT SIGNAL 1 21
 FT CHAIN 1 22
 FT DISULFID 28 326
 FT DISULFID 50 172
 FT CARBOHYD 197 197
 FT CARBOHYD 206 206
 FT CARBOHYD 235 235
 FT CARBOHYD 249 249
 FT CARBOHYD 256 256
 SO SEQUENCE 326 AA; 34556 MW; F99D7E77F896FA2C CRC64;

OY 12 L1LSTLSPLGLPVLCAPPLICDSRVLEKRYLLEAKENITTCGAEHCSLNENITVPDT 71

OY 12 L1LSTLSPLGLPVLCAPPLICDSRVLEKRYLLEAKENITTCGAEHCSLNENITVPDT 71

Query Match 9.0%; Score 90; DB 1; Length 356;
 Best Local Similarity 25.1%; Pred. No. 0.39; Mismatches 78; Indels 24; Gaps 6;
 Matches 44; Conservative 29;

Query Match 8.7%; Score 86.5; DB 1; Length 326;
 Best Local Similarity 23.0%; Pred. No. 0.75; Mismatches 73; Indels 41; Gaps 6;
 Matches 42; Conservative 27;


```

Db 7 LVAALLTLARL-FLSSVPAPCPRLNKLRLSDYLSHLSLSCPDVNPISIPVLFVAV 65
Qy 72 KVFNFYAKRMEVROQAVEVWOGIALLSEAVL-RGQALLVNSOPWEPLQIHDVKAASGL 129
Db 66 DFLSGEKKKTQTESKADIDIGAVSLLEGVMAARGQL-----EF----- 104
Qy 130 RSLTTLRALGAQKEAI-----SPDDAASAPLRTITADTFERKLFVYSNFLRG 178
Db 105 SCLSSLLGQLSGQVRLILGALGILGTLPLPQG-----FTTAKHDPALFLSLQQLLRG 158
Qy 179 KTK 181
Db 159 KVR 161

RESULT 15
FACG_HUMAN
ID FACG_HUMAN STANDARD: PRT: 622 AA.
AC 015287;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fanconi anemia group G protein (FACG protein) (DNA-repair protein
DE XRC9).
GN FACG OR XRC9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97404378; PubMed=9256465;
RA Liu N., Lamerdin J.E., Tucker J.D., Zhou Z.-Q., Walter C.A.,
RA Albata J.S., Busch D.B., Thompson L.H.;
RT "The human XRC9 gene corrects chromosomal instability and mutagen
RT sensitivities in CHO UV40 cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:9232-9237(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99021385; PubMed=9806548;
RA De Winter J.P., Walsflisz Q., Koolmans M.A., Van Berkel C.G.M.,
RA Bosnyan Collins L., Alon N., Carreau M., Bender O., Demuth I.,
RA Schindler D., Pronk J.C., Arwert F., Hoehn H., Digweed M.,
RA Buchwald M., Joenje H.;
RT "The Fanconi anaemia group G gene FANCG is identical with XRC9.";
RL Nat. Genet. 20:281-283(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Fibroblast, and Foreskin;
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Adamson A.W.,
RA Burkhart-Schultz K., Gordon L., Kyle A., Ramirez M., Stillwagen S.,
RA Phan H., Velasco N., Ganes J., Dangman L., Poundstone P.,
RA Christensen M., Georgescu A., Avila J., Liu S., Attik C., Andreise T.,
RA Tranhhe M., Amico-Keller G., Coetfield J., Duarte S., Lucas S.,
RA Bruce R., Thomas P., Quan G., Krommiller B., Arellano A.,
RA Montgomery M., Ow D., Nolan M., Truong S., Kobayashi A., Olsen A.O.,
RA Carrano A.V.;
RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
RN [5]
RP CHARACTERIZATION.
RX MEDLINE=99303779; PubMed=10373536;
RA Garcia-Higuera I., Kuang Y., Nat D., Wasik J., D'Andrea A.D.;
RT "Fanconi anemia proteins FANCA, FANCC, and FANCG/XRC9 interact in a
RT functional nuclear complex.";
RL Mol. Cell. Biol. 19:4866-4873(1999).
CC -1- FUNCTION: DNA REPAIR PROTEIN THAT MAY OPERATE IN A POSTREPLICATION
CC REPAIR OR A CELL CYCLE CHECKPOINT FUNCTION. MAY BE IMPLICATED IN

```

```

CC INTERSTRAND DNA CROSS-LINK REPAIR AND IN THE MAINTENANCE OF NORMAL
CC CHROMOSOME STABILITY. CANDIDATE TUMOR SUPPRESSOR GENE.
CC -1- SUBUNIT: BELONGS TO A MULTISUBUNIT COMPLEX COMPOSED OF FANCA,
CC FANCB AND FANCG. THE COMPLEX IS NOT FOUND IN PA PATIENTS.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (MAJOR) AND CYTOPLASMIC (MINOR).
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTIS AND THYMOS. FOUND
CC IN LYMPHOBLASTS.
CC -1- DISEASE: DEFECTS IN FANCG ARE THE CAUSE OF ONE OF THE 8
CC COMPLEMENTATION GROUPS OF FANCONI ANEMIA (FA), AN AUTOSOMAL
CC RECESSIVE DISORDER CHARACTERIZED BY PROGRESSIVE PANCYTOPEANIA, A
CC DIVERSE ASSORTMENT OF CONGENITAL MALFORMATIONS, AND A
CC PREDISPOSITION TO THE DEVELOPMENT OF MALIGNANCIES. AT THE CELLULAR
CC LEVEL, IT IS ASSOCIATED WITH HYPERSENSITIVITY TO DNA-DAMAGING
CC AGENTS, CHROMOSOMAL INSTABILITY (INCREASED CHROMOSOME BREAKAGE),
CC AND DEFECTIVE DNA REPAIR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U70310; AAB80802.1; -
DR EMBL: AJ007669; CAA07602.1; -
DR EMBL: AC004472; AAC07981.1; -
DR EMBL: BC000032; AAH00032.1; -
DR Genew: HGNC:3588; FANCG.
DR MIM: 602956; -
DR InterPro: IPR01440; TPR.
DR Pfam: PF00515; TPR; 2.
KW DNA repair; Nuclear protein.
SQ SEQUENCE 622 AA; 68553 MW; 4BC7475472AC3C84 CRC64;

Query Match 8.7%; Score 86.5; DB 1; Length 622;
Best Local Similarity 26.8%; Pred. No. 1.7;
Matches 51; Conservative 23; Mismatches 81; Indels 35; Gaps 7;

Qy 12 LLSLSLPLGLPVLCAPRLICDSRYLRLLEKKAENTTGCAGHC--SLNENITVP 69
Db 53 LHSIQLPAAVPLPLEITYTCNF-----ILRASLQGTEDQADQISLSLR--VL 104
Qy 70 DTKVNFYAKRMEVROQAVEVWOGIALLSEAVLRGQAL--LVNSOPWEPLQIHDVKA 126
Db 105 ETQ-----EQGPRLEQGLRELMDSVLRASCLLPFLSALHLYVGLQAALWLSA 153
Qy 127 SGLSLTTLRAL-----GAQKEAI-----SPDDAASAPLRTITADTFERKLFVYSN 175
Db 154 DRLGDLALLEFTLNGSGGASKDLILLIKTWSPEEIDADLTLDAGLKDVLTAFAV 213
Qy 176 LRGLKLKLYTG 185
Db 214 RGLQELLYTG 223

```

Search completed: January 8, 2003, 03:43:19
 Job time : 23.4401 secs


```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_taxid=9986;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=21290682; PubMed=11396976;
RA Vialata A., Wu D., Margalich M., Hobart P.;
RT "Rabbit EPO Gene and cDNA: Expression of Rabbit EPO after
RL Intramuscular Injection of pDNA".
RL Biochem. Biophys. Res. Commun. 284:823-827(2001).
DR EMBL: AF280944; AAG36962.1; -.
DR HSSP; P01588; ICN4.
DR InterPro: IPR001323; EPO_TPO.
DR InterPro: IPR003013; Erythropn.
DR Pfam: PF00758; EPO_TPO.1.
DR PRINTS; PR00272; ERYTHROPTN.
DR PROSITE; PS00817; EPO_TPO.1.
SQ SEQUENCE 195 AA; 21025 MW; 1F1DC7F403A303EC CRC64;

Query Match 73.0%; Score 728.5; DB 6; Length 195;
Best Local Similarity 80.8%; Pred. No. 9,2e-64;
Matches 147; Conservative 12; Mismatches 22; Indels 1; Gaps 1.

QY 13 LLSLSLPLGLPVLCAPPRLICDSVLERYLEAKAEENITTCAGHCSSLNENITVPDTR 72
Db 14 LLLCLLVALLGLPVLCAPPRLICDSVLERYLEAKAEENITTCAGHCSSLNENITVPDTR 73
QY 73 VNFYAMKMEYRROQAVEWOGALLSEVLRGQALLVSSQGWPELQCHVDKAVSGIRSL 132
Db 74 VNFHMKSESGRAHAEVWQGLALSEMLRSQALLANSQPELQYHVDKAVSGIRSL 133
QY 133 TTLRLALGAKREKASPDDA-SAAPLRTITADTFPKLFRVYSNPLRGKLTGTGACRTG 191
Db 134 TSLRLALGVQKEAVSPPAASADPLRFVADTLCKLFRYSNPLRGKLTGTGACRNG 193
QY 192 DR 193
Db 194 DR 195

RESULT 4
O9QV40 PRELIMINARY; PRT; 50 AA.
AC O9QV40;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Erythropoietin (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_taxid=10118;
RN
RP SEQUENCE.
RX MEDLINE=94115047; PubMed=7764337;
RA Okano M., Suga H., Masuda S., Nagao M., Narita H., Ikura K.,
RL Sasaki R.;
RL Biosci. Biotechnol. Biochem. 57:1882-1885(1993).
RL HSSP; P01588; IEBR.
DR InterPro: IPR001323; EPO_TPO.
DR InterPro: IPR003013; Erythropn.
DR Pfam: PF00758; EPO_TPO.1.
DR PRINTS; PR00272; ERYTHROPTN.
DR SEQUENCE 50 AA; 5587 MW; 70B44ABBF016034 CRC64;

Query Match 18.8%; Score 188; DB 11; Length 50;
Best Local Similarity 78.0%; Pred. No. 2,6e-11;
Matches 39; Conservative 3; Mismatches 8; Indels 0; Gaps 0.

QY 28 APPRLICDSVLERYLEAKAEENITTCAGHCSSLNENITVPDTRKVNFYA 77
Db 1 APPRLICDSVLERYLEAKAEENITTCAGHCSSLNENITVPDTRKVNFYA 50

```

```

RESULT 5
O9RPH5 ID O9RPH5 PRELIMINARY; PRT; 554 AA.
AC O9RPH5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE RECD.
GN RECD.
OS Mycobacterium smegmatis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1772;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC2155;
RX MEDLINE=99412429; PubMed=10481025;
RA Griffin IV T.J., Parsons L., Leschziner A.E., DeVost J.,
RA Derbyshire K.M., Grindley N.D.;
RT "In vitro transposition of tm552: a tool for DNA sequencing and
RT mutagenesis."
RL Nucleic Acids Res. 27:3859-3865(1999).
DR EMBL: AF157643; AAD46809.1;
DR InterPro: IPR000606; Viral_helicase1.
DR Pfam: PF01443; Viral_helicase1.1;
SQ SEQUENCE 554 AA; 59516 MW; 3947B50498BD62C CRC64;

Query Match 9.5%; Score 94.5; DB 2; Length 554;
Best Local Similarity 26.7%; Pred. No. 0.89; Mismatches 56; Indels 61; Gaps 9;
Matches 50; Conservative 20;

OY 8 AMWLILSLPLGLPVLGAP--RLICDSRVLEKLEAKENITTCGAEHCSTL--- 62
DB 72 AMLAALAA-----SPILGQPVRLRFGDLTYLDRIYWLAEQY-----CDDVIALVSA 118
OY 63 NENTIVPDTKYNFYAMKMEYRQQAWEYWGCLALSLSE-----AVLRGQALL 108
DB 119 RFGCAVDVSRFLGFCAGEEORAAKVAISGLITVLTGPGTGKTTVARRLLALAEQAL 178
OY 109 VNSSQP-----WEPLQLQVHD-----KAVSGLSLTTLRLALGAKRAI 146
DB 179 AGKSPRIALAAPTGKAAARLQEAIVQLEIDLDLIRRRRLGLHA--TTLHRLGPR----- 233
OY 147 SPDDAAS 153
DB 234 --PDTSS 238

RESULT 6
O9ROR6 ID O9ROR6 PRELIMINARY; PRT; 623 AA.
AC O9ROR6;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Fanconi anemia group G protein (Fanconi anemia complementation group
DE 6).
GN FANCG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB;
RA Tipping A.J., Morgan N.V., Mathew C.G.;
RT "Sequencing and genomic structure of fangc, the murine orthologue of
RT the Fanconi anemia group G gene (FANCG).";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB;
RA van de Vugt H.J., Koomen M., Berns M.A.B., de Vries Y.,

```

```

RA Roimans M.A., van der Weel L., Blom E., de Winter J.P., de Groot J.,
RA Schepers R.J., Hoatlin M.E., Ching Cheng N., Joenje H., Arwert F.;
RT "Characterization, expression and complex formation of murine Fancg.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF112439; AAC63198.1;
DR MGD: MGI:1926471; Fancg.
DR InterPro: IPR001440; TPR.
DR Pfam: PF00515; TPR.
DR SMART: SM00028; TPR.2;
SQ SEQUENCE 623 AA; 68505 MW; 061586EF186F74AF CRC64;

Query Match 9.4%; Score 94; DB 11; Length 623;
Best Local Similarity 25.1%; Pred. No. 1.2;
Matches 51; Conservative 25; Mismatches 65; Indels 62; Gaps 8;

```

```

OY 12 LLSLSLPLGLPVLGAPPRILICDSRVLEKLEAKENITTCG-----AEHCSTL 62
DB 58 LLTLTGPAVPAVPALEPLTLVNCITLRASLYQA-FTEEDTLQDLQGLERVENQH--- 113
OY 63 NENTIVPDTKYNFYAMKMEYRQQAWEYWGCLALSLSEAVLRGQAL--LVNSSQWPEPLQ 119
DB 114 -----HLERKSQ-----QGLKELMHVSLASSLPPELLALHCLASLQ 151
OY 120 LHVDAVSGLSLTTLRLAL-GAOKRAI-----SPDDASADPLRTTADTFKRL 168
DB 152 AVFWMTDHLDELTLTLQTLNGSCFOSSEDLILLKSWSPAESPADLIQDASLRDY 211
OY 169 FRVYSNPLRGLKLYTGCACTRG 191
DB 212 -----LTTAPACRQG 221

```

```

RESULT 7
O8ZKZ4 ID O8ZKZ4 PRELIMINARY; PRT; 346 AA.
AC O8ZKZ4;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Periplasmic sensor in multi-component regulatory system with Tors
DE (sensory kinase) and TorC (regulator), regulates for operon.
GN TORC OR STM3825.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Portwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856(2001).
DR EMBL: AE008878; AAL22684.1;
DR InterPro: IPR001761; Periplasm/LacI.
DR Pfam: PF00532; Peripla_BP_1like; 1.
KW Complete proteome.
SQ SEQUENCE 346 AA; 38440 MW; E37CAB58E49FD716 CRC64;

```

```

Query Match 8.8%; Score 87.5; DB 16; Length 346;
Best Local Similarity 26.7%; Pred. No. 2.4;
Matches 44; Conservative 22; Mismatches 48; Indels 51; Gaps 9;

OY 37 RVLEKYLEAKENITTCG--CAEHCSTLNE--NITVPTKYNFYAMKMEYRQQAWEYWG 92
DB 217 RNLIQEMLERHPADVAVAGSAIAAEAMGEGRNLTPLTIVSYL-----THOYR 267

```

```

QY 93 GLALLSEAVLRGQALLVNSQ--PWEPLQLHVDKAVSGLSLTLRALGAQ--KEAISP 149
DB 268 GLK-----RGHILMALSDQAMQ-----GELAITQSIKVLQGPVPEINISPP 309
QY 150 -----DASAPLRTITADTFPRKLFRRYSNLRGKLIKLYTGEA 187
DB 310 VLITLHNNDASARVRSLSPPGFRPVY-----LYQYTSSEA 344

RESULT 8
082ZM5 PRELIMINARY: PRT: 346 AA.
AC 082ZM5:
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Solite binding receptor protein.
GN STY3952.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxId=601;

RN SEQUENCE FROM N.A.
RC STRAIN=CT18:
RA MEDLINE=21534947; PubMed=11677608;
RA Parhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Croft A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18."
RL Nature 413:848-852(2001).
DR EMBL: AL627280; CAD03169.1;
DR InterPro: IPR001761; PeriplabP/LacI.
DR Pfam: PF00532; Peripla_BP_1like; 1.
KW Receptor; Complete Proteome.
SQ SEQUENCE 346 AA; 38546 MW; F80FBL688BC83A8F CRC64;

Query Match 8.8%; Score 87.5; DB 16; Length 346;
Best Local Similarity 26.7%; Pred.No.2.4;
Matches 44; Conservative 22; Mismatches 48; Indels 51; Gaps 9;

QY 37 RYLERYLTAKEAENTTTG--CAEHCSLNE--NITVPDRKVNFAKRMREVRQOAVEWQ 92
DB 217 RNLQEMLEHRHDANVYAGSAIAEAAGGRULTPRTIVSYFL-----THQYR 267
QY 93 GLALLSEAVLRGQALLVNSQ--PWEPLQLHVDKAVSGLSLTLRALGAQ--KEAISP 149
DB 268 GLK-----RGHILMALSDQAMQ-----GELAITQSIKVLQGPVPEINISPP 309
QY 150 -----DASAPLRTITADTFPRKLFRRYSNLRGKLIKLYTGEA 187
DB 310 VLITLHNNDASARVRSLSPPGFRPVY-----LYQYTSSEA 344

RESULT 9
09D7X0 PRELIMINARY: PRT: 454 AA.
AC 09D7X0:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE 2210018A16Rik protein.
GN ZFP296 OR 2210018A16Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=STOMACH;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi T., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Koehli H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boileau D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazarrelli J., Momberts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,
RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL: AK008746; BAB25873.1;
DR HSSP: P08153; 1ZFD.
DR MGI: 1926956; Zfp296.
DR InterPro: IPR000822; Znf.C2H2.
DR Pfam: PF00096; zfp.C2H2; 5.
DR SMART: SM00355; Znf.C2H2; 5.
DR PROSITE: PS00028; ZINC_FINGER_C2H2.1; 4.
DR PROSITE: PS00157; ZINC_FINGER_C2H2.2; 4.
DR DNA-binding; Metal-binding; zinc-finger.
SQ SEQUENCE 454 AA; 48934 MW; 9012C24C9AB02000 CRC64;

Query Match 8.8%; Score 87.5; DB 11; Length 454;
Best Local Similarity 24.0%; Pred.No.3.4;
Matches 48; Conservative 24; Mismatches 81; Indels 47; Gaps 8;

QY 14 LSLSLPLGLPVLAGAPRLICDSRYLERYLTAKEAENTTTG--CAEHCSLNE--NITVPDRK 72
DB 58 LQFASRPLAPSTCA--PRNPPLSKSSDRQPMWDKHHDLTLGCGKGFPLGATIANMDK 116
QY 73 VNFYAKRMREVRQOAVEWQGLAL--SEAVLRGQALLVNS-----SQWEPI----- 118
DB 117 K-----QGCQLQVSPDISSEKELKAPSCLOCGROYTSPMKILCHAO 158
QY 119 -----QLHYD--KAVSGRLSTLTLRALGAQKEAISPDAAPLRTITADTFR 166
DB 159 WDHGCLTYOTQHLDPTEAPLGLAEVAAASNAVAVPESKPPVSSAARSPTCDVCK 218
QY 167 KLFRRYSNLRGKLIKLYTGE 186
DB 219 KTLSSFSN-LKVMRSHGTGE 237

RESULT 10
08ZDC8 PRELIMINARY: PRT: 323 AA.
AC 08ZDC8:
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Ribonucleoside-diphosphate reductase 2 beta chain (EC 1.17.4.1).
GN NRDP OR YPO2648.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxId=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92; BIOVAR ORIENTALIS;
RX MEDLINE=21470413; PubMed=11586360;

```

RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
 RA Prentice M.B., Sebaitia M., James K.D., Churcher C., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarrega A.M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Felwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
 RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
 RA Simmonds M., Skellon J., Stevens K., Whitehead S., Barrell B.G.,
 RT "genome sequence of Yersinia pestis, the causative agent of plague";
 RL Nature 413:523-527(2001).
 DR EMBL: AJ414153; CAC92889.1; -
 DR InterPro: IPR000358; RibonucL_redctase.
 DR Pfam: PF00268; ribonuc_red_sm: 1.
 DR PROSITE: PS00368; RIBOPED_SMALL: 1.
 KW Oxidoreductase; Complete proteome.
 SQ SEQUENCE 323 AA; 36823 MW; 87C21F7BB9B7FD2 CRC64;
 Query Match 8.7%; Score 87; DB 16; Length 323;
 Best Local Similarity 25.2%; Pred. No. 2.4;
 Matches 34; Conservative 20; Mismatches 59; Indels 22; Gaps 5;
 QY 65 NITVPDTKVNPFYAMKRMVEVROQAVEVWQGLALLSEAVLRGALLVNSQPEPQLQHYD- 123
 DB 2 NVKPIRIRISAINMKIE-DDKLELVNN--RLTSNFWLPKXVPLSNDIPSWATLTPHROQ 58
 QY 124 ---KAVSGLSLTLLALANQ---KEAISPPDAASAPLRTITADTFKLFYRYSNFLR 177
 DB 59 LTRVFTGLTLDTLQNTLQAPALIKDAIIPHEAIFISIFMEAVHARSYSIFSTL-- 116
 QY 178 GKLLTYGECACRTGD 192
 DB 117 -----CLTSD 121
 RESULT 11
 Q8T527 PRELIMINARY; PRT; 877 AA.
 AC Q8T527;
 DT 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DE ABC transporter ABCA.1.
 DE ABC transporter ABCA.1.
 GN ABCA.1.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AX4;
 RA Anjard C., Loomis W.F.,
 RT "Evolution of the ABC transporter of Dictyostelium,"
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF491005; AAL99041.1; -
 SQ SEQUENCE 877 AA; 97375 MW; DBEC7C7ECB34309 CRC64;
 Query Match 8.7%; Score 87; DB 5; Length 877;
 Best Local Similarity 27.2%; Pred. No. 8.7;
 Matches 28; Conservative 9; Mismatches 44; Indels 22; Gaps 2;
 QY 14 LSLSLPLGPIVCAAPR-----LTCDSRVLERYLLEAKENITWGAEHSL----- 62
 DB 170 VDFLSTPLGQINGCAPKNEPLKMDGNSLEFNYYVNTTCNIIATCPDIYIAITTAIEK 229
 QY 63 -----NENITVPDTKVNPFYAMKRMVEVROQAVEVWQGL 94
 DB 230 AVITVYSQIRNEKIPPTISYGSNAPRYPRPEGGAARVWGGL 272
 RESULT 12
 Q9HZM7 PRELIMINARY; PRT; 339 AA.
 AC Q9HZM7;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)

DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE UDP-N-acetylpyruvoylglycosamine reductase.
 GN MURB OR PA2977.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.V., Lagrou M.,
 RA Gader R.L., Gollery L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 opportunistic pathogen";
 RL Nature 406:959-964(2000).
 DR EMBL: AE004723; AAC06365.1; -
 DR HSSP: P08373; 2MBR.
 DR InterPro: IPR003170; MurB.
 DR InterPro: IPR001575; Oxid_FAD_bind.
 DR InterPro: IPR000531; TonB_boxC.
 DR Pfam: PF01565; FAD_binding_4; 1.
 DR Pfam: PF02873; MurB_C; 1.
 DR PROSITE: PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 339 AA; 37627 MW; C0C8EF9P2938FE27 CRC64;
 Query Match 8.5%; Score 85; DB 16; Length 339;
 Best Local Similarity 23.0%; Pred. No. 4.1;
 Matches 42; Conservative 22; Mismatches 49; Indels 70; Gaps 8;
 QY 58 EHCSINE-NITVPDTKVNPFYAMKRMVEVROQAVEVWQGLALLSEAVLRGALLV----- 109
 DB 7 EHCSIKPYTEGIVRARLAHARDE-----ADVREALALARE--RLPLVLVIGGSNL 58
 QY 110 -----NSQPEP-IQLHYDAVSGLSLTLL 136
 DB 59 LTRDVEALVLRMASQGRIVSDAADSVLVEAEAGEMDPVQNSLEGLAGLESLI- 117
 QY 137 RALGAQKEAISPPDAASAPLRTITA-----DFFKLFYRYSNFLRKLTYGECART 190
 DB 118 -----PGTVGAAPQINIGAYGVLEKIDVDSLTAI--DRODCTLREFDQACRF 163
 QY 191 GDR 193
 DB 164 GYR 166
 RESULT 13
 Q8ZHV3 PRELIMINARY; PRT; 2201 AA.
 AC Q8ZHV3;
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DE Putative siderophore biosynthesis protein.
 GN YP00778.
 OS Yersinia pestis.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Yersinia.
 OX NCBI_TaxID=632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CO-92 / BIOVAR ORIENTALIS;
 RX MEDLINE=21470413; PubMed=11586360;
 RA Parkhill J., Wren B.W., Thomson N.R., Churcher C., Mungall K.L.,
 RA Prentice M.B., Sebaitia M., James K.D., Churcher C., Cerdano-Tarrega A.M.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Davis P., Dougan G.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,

RA Felwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
 RA Leathers S., Moule S., Oyston P.C.F., Unali M., Rutherford K.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.,
 RT "Genome sequence of Yersinia pestis, the causative agent of plague";
 RL Nature 413:523-527(2001).
 DR EMBL: AJ414144; CAC89627.1; -
 DR InterPro: IPR000873; AMP-bind.
 DR InterPro: IPR001242; Condensatin.
 DR InterPro: IPR003880; Ppantne_attach.
 DR Pfam: PF00501; AMP-binding; 1.
 DR Pfam: PF00668; Condensation; 2.
 DR Pfam: PF00550; pp-binding; 2.
 DR PRINTS: PR00154; AMPBINDING.
 DR PROSITE: PS0075; ACP_DOMAIN; 2.
 DR PROSITE: PS00455; AMP_BINDING; 1.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
 KW Hypothetical protein: Complete proteome.
 SQ SEQUENCE 2201 AA; 246655 MW; 168E22019020809B CRC64;

Query Match 8.5%; Score 84.5; DB 16; Length 2201;
 Best Local Similarity 25.5%; Pred. No. 49;
 Matches 48; Conservative 18; Mismatches 55; Indels 67; Gaps 9;

QY 23 LPVLGAPPR-LICDSRYLE-----RYLLEAKENITTC----- 56
 DB 1910 LPMPLPAPRLPLVCDPALVGTPEKFKMOAVLASDMQMLHKAKHQITPSCILLTCYAO 1169.
 QY 57 -----AEHCSLSENIITVPPTKV-----NFYAMKREYVROQAVEVWOGIALISEAV 101
 DB 1970 VLAKMESASLTINTVTLFDRKPCHPDINHMGFTSLILGCGEKGEGMLATA----- 2023
 QY 102 LKGOALLVNSQPEPLQLHVDKAVSGLSL-----TTLIR-----ALCAQKEAISP 149
 DB 2024 ---QRI---QOQLMRDLREHREVSAAVWLRELNRQGTTHIMPEVETSALGAQVDGY--P 2075
 QY 150 DASAAPL 157
 DB 2076 DSAFQAPL 2083

RESULT 14
 ID P94873 PRELIMINARY; PRT: 3722 AA.
 AC P94873:
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-MAY-2002 (TREMBLrel. 20, Last annotation update)
 DE Alpha-aminoadipyl-cysteineyl-valine synthetase.
 GN PCBA8.
 OS Bacteroides lactamgenus.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Lysobacter.
 OX NCBI_TaxID=39596;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-YK90;
 RX MEDLINE=96275949; PubMed=8737573;
 RA Kimura H., Miyashita H., Sumino Y.,
 RT "Organization and expression in Pseudomonas putida of the gene cluster
 RT involved in cephalosporin biosynthesis from lysobacter lactamgenus
 RT YK90";
 RL Appl. Microbiol. Biotechnol. 45:490-501(1996).
 DR EMBL: D50308; BAA08846.1; -
 DR HSP: P14687; IAMU.
 DR InterPro: IPR002106; ALRNA_ligaseII.
 DR InterPro: IPR000873; AMP-bind.
 DR InterPro: IPR001242; Condensatin.
 DR InterPro: IPR000977; DNA_ligase.
 DR InterPro: IPR003880; Ppantne_attach.
 DR InterPro: IPR000379; Ser_estrs_site.
 DR InterPro: IPR001031; Thioesterase.
 DR Pfam: PF00501; AMP-binding; 3.
 DR Pfam: PF00668; Condensation; 3.

DR Pfam: PF00550; pp-binding; 3.
 DR Pfam: PF00975; Thioesterase; 1.
 DR PROSITE: PS00179; AA_TRNA_LIGASE_II_1; UNKNOWN_1.
 DR PROSITE: PS0075; ACP_DOMAIN; 3.
 DR PROSITE: PS00455; AMP_BINDING; 1.
 DR PROSITE: PS00697; DNA_LIGASE_AI; UNKNOWN_2.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; UNKNOWN_2.
 KW Phosphopantetheine.
 SQ SEQUENCE 3722 AA; 411607 MW; 3597B3483463809B CRC64;

Query Match 8.5%; Score 84.5; DB 2; Length 3722;
 Best Local Similarity 20.1%; Pred. No. 97;
 Matches 39; Conservative 20; Mismatches 60; Indels 75; Gaps 7;

QY 4 HEC-PAMWLLSLPLGLPVLGAPPRLLICDSRYLERYLEAKENITTCGACRCSL 62
 DB 2120 HSCFDKSMKIF-----RELQALLGATAGTLP----- 2148
 QY 63 NENITVPDTIKNFYAMKREYVROQAVEV-----WOGIALISEAVLRCOALLVNSQPEPL 118
 DB 2149 ---PLPATVADPSVWQROOLSDQRLDALPDYWQ-----RSLAGMPL 2187
 QY 119 QLVHDKAVSGLSRLTTLRLAGNO-----KEAISPDAASAAPLRTITADTFKKLRYS 173
 DB 2188 QLPD-----HARPAQFDYIGREIVFDVADATTCQDLRVLAQTTSTFSVLL 2234
 QY 174 NFLRGRLKLYTGEA 187
 DB 2235 AAYVTLTKAVSGQS 2248

RESULT 15
 ID Q9QVR7 PRELIMINARY; PRT: 235 AA.
 AC Q9QVR7:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Truncated thrombopoietin.
 DE Mus sp.
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10095;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95382808; PubMed=7654225;
 RA Wada T., Nagata Y., Nagahisa H., Okutomi K., Ha S.H., Ohnuki T.,
 RA Kanaya T., Matsumura M., Todokoro K.,
 RT "Characterization of the truncated thrombopoietin variants";
 RT Biochem. Biophys. Res. Commun. 213:1091-1098(1995).
 RL InterPro: IPR001323; EPO_TPO.
 DR InterPro: IPR003978; thrombopoietin.
 DR Pfam: PF00758; EPO_TPO; 1.
 DR PRINTS: PR01485; THROMBOPTN.
 DR PROSITE: PS00817; EPO_TPO; 1.
 SQ SEQUENCE 235 AA; 25810 MW; DABA8DC2158F7C91 CRC64;

Query Match 8.4%; Score 84; DB 11; Length 235;
 Best Local Similarity 25.0%; Pred. No. 3.2;
 Matches 24; Conservative 22; Mismatches 48; Indels 2; Gaps 1;

QY 12 LLLSLSLPLGLPVLGAPPRLLICDSRYLERYLEAKENITTCGACRCSLSENIITVPPT 71
 DB 6 LLLAAMLAVALRLLTSSPAACDPRLKLRLDLSRLSLSGCPDVPDPLSPVLLP 65
 QY 72 KVFYAMKREYVROQAVEVWOGIALISEAVL--RCQ 105
 DB 66 DPLSGEMKTQTEQSKADIIIGAVSLLLEGVMAARQ 101

Search completed: January 8, 2003, 05:13:46
 Job time : 74.873 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 03:39:48 ; Search time 12.9025 Seconds
(without alignments)
440.118 Million cell updates/sec

Title: US-09-813-775C-2

Perfect score: 998
Sequence: 1 MGVECPAMWLILSLSLP.....NPLRGLKLYTGACRTGDR 193

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued_Patents_AA:*
2: /cgn2_6/prodata1/1aa/5a_COMB.pep:*
3: /cgn2_6/prodata1/1aa/5b_COMB.pep:*
4: /cgn2_6/prodata1/1aa/6a_COMB.pep:*
5: /cgn2_6/prodata1/1aa/6b_COMB.pep:*
6: /cgn2_6/prodata1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	99.1	99.3	193	1	US-07-903-220-1
2	99.1	99.3	193	2	US-08-883-795A-34
3	84.6	84.8	412	4	US-09-366-009-34
4	84.3	84.5	166	1	US-08-318-193-70
5	84.3	84.5	166	4	US-09-604-871-2
6	84.3	84.5	166	5	PCT-US94-04361-37
7	83.8	84.0	165	4	US-09-604-871-1
8	83.5	83.7	166	5	PCT-US94-04361-45
9	76.1	76.3	165	5	PCT-US94-04361-38
10	75.6	75.8	165	5	PCT-US94-04361-39
11	71.0	71.1	166	5	PCT-US94-04361-44
12	70.8	70.9	166	5	PCT-US94-04361-41
13	69.6	69.7	166	5	PCT-US94-04361-40
14	68.5	68.8	167	5	PCT-US94-04361-42
15	68.6	68.7	168	5	PCT-US94-04361-43
16	18.5	18.5	36	5	PCT-US94-04361-50
17	16.6	16.6	30	5	PCT-US94-04361-28
18	15.6	15.6	30	5	PCT-US94-04361-29
19	14.8	14.8	426	3	US-08-932-823A-2
20	12.8	12.8	27	5	PCT-US94-04361-51
21	11.4	11.5	34	5	PCT-US94-04361-52
22	10.8	10.8	20	2	US-08-759-599-9
23	10.8	10.8	20	4	US-09-294-457-9
24	10.7	10.7	21	5	PCT-US94-04361-49
25	10.6	10.6	30	5	PCT-US94-04361-31
26	10.3	10.3	25	5	PCT-US94-04361-34
27	10.1	10.2	25	5	PCT-US94-04361-32

28	10.1	10.1	20	5	PCT-US94-04361-58	Sequence 58, Appl
29	98.5	9.9	25	5	PCT-US94-04361-35	Sequence 35, Appl
30	96.5	9.7	174	3	US-08-471-045-56	Sequence 56, Appl
31	96.5	9.7	174	3	US-08-469-712A-56	Sequence 56, Appl
32	96.5	9.7	174	4	US-08-446-871-56	Sequence 56, Appl
33	96.5	9.7	174	4	US-08-468-910-56	Sequence 56, Appl
34	96.5	9.7	174	4	US-08-761-907-56	Sequence 56, Appl
35	96.5	9.7	195	1	US-08-388-779A-4	Sequence 56, Appl
36	96.5	9.7	195	1	US-08-388-779A-4	Sequence 4, Appl
37	96.5	9.7	195	1	US-08-927-855-4	Sequence 4, Appl
38	96.5	9.7	353	1	US-08-330-517-2	Sequence 2, Appl
39	96.5	9.7	353	1	US-08-347-029-4	Sequence 2, Appl
40	96.5	9.7	353	1	US-08-388-779A-2	Sequence 2, Appl
41	96.5	9.7	353	1	US-08-484-246-1	Sequence 1, Appl
42	96.5	9.7	353	1	US-08-591-070A-2	Sequence 2, Appl
43	96.5	9.7	353	1	US-08-413-803-25	Sequence 25, Appl
44	96.5	9.7	353	1	US-08-321-488A-25	Sequence 25, Appl
45	96.5	9.7	353	2	US-08-414-161B-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-07-903-220-1
Sequence 1, Application US/07903220
Patent No. 5322837
GENERAL INFORMATION:
APPLICANT: Hewlett, Rodney M.
TITLE OF INVENTION: METHOD FOR THE PURIFICATION OF
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Paul H. Heller
STREET: Kenyon & Kenyon, One Broadway
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/903,220
FILING DATE: 19920731
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: 1248/27
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 429-1776
TELEFAX: (202) 429-0796
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 193 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-07-903-220-1

Query Match 99.3%; Score 991; DB 1; Length 193;

Best Local Similarity 99.5%; Pred. No. 2.9e-111;
Matches 192; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MGVECPAMWLILSLSLPGLGVLCAPPRLICDSRYLEKYLEAKAEKAITTGCAEHC 60
DB 1 MGVECPAMWLILSLSLPGLGVLCAPPRLICDSRYLEKYLEAKAEKAITTGCAEHC 60

OY 61 SLNENITVDTKVNFAKRMVEVROQAVEWOGALLSEAVLRGALLVNSSQPMPEPLOL 120
DB 61 SLNENITVDTKVNFAKRMVEVROQAVEWOGALLSEAVLRGALLVNSSQPMPEPLOL 120
OY 121 HYDKAVSGRLSTLTLLRALGAQKEAISPPDASAAPLRTITADTFKRLFRVYSNPLRGKL 180
DB 121 HYDKAVSGRLSTLTLLRALGAQKEAISPPDASAAPLRTITADTFKRLFRVYSNPLRGKL 180
OY 181 KLYTGEACRTGDR 193
DB 181 KLYTGEACRTGDR 193

RESULT 2
US-08-883-795A-34
Sequence 34, Application US/08883795A
Patent No. 5985607
GENERAL INFORMATION:
APPLICANT: Delcuve, Genevieve
APPLICANT: Awang, Gregor
TITLE OF INVENTION: Recombinant DNA Molecules and Expression
TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: BERESKIN & PARR
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/883,795A
FILING DATE: 27-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gravelle, Micheline
REGISTRATION NUMBER: 40,261
REFERENCE/DOCKET NUMBER: 7841-062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 193 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-883-795A-34

Query Match 99.3%; Score 991; DB 2; Length 193;
Best Local Similarity 99.5%; Pred. No. 2.9e-111;
Matches 192; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 181 KLYTGEACRTGDR 193

RESULT 3
US-09-366-009-34
Sequence 34, Application US/09366009
Patent No. 6426042
GENERAL INFORMATION:
APPLICANT: Asada, Kiyozo
Uemori, Takashi
Ueno, Takashi
Koyama, No. 6426042uto
Hashino, Kimikazu
Kato, Ikunoshin
TITLE OF INVENTION: METHOD FOR GENE TRANSFER INTO TARGET
CELLS WITH RETROVIRUS
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: WEISER & ASSOCIATES
STREET: 230 South Fifteenth Street, Suite 500
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19102
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/366,009
FILING DATE: 02-Aug-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/809,156
FILING DATE: 13-NOV-1995
APPLICATION NUMBER: JP 294382/1995
FILING DATE: 13-NOV-1995
APPLICATION NUMBER: JP 051847/1996
FILING DATE: 08-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Weiser, Gerard J.
REGISTRATION NUMBER: 19,763
REFERENCE/DOCKET NUMBER: 977.6507P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-875-8383
TELEFAX: 215-875-8394
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 412 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-09-366-009-34

Query Match 84.8%; Score 846; DB 4; Length 412;
Best Local Similarity 98.8%; Pred. No. 2.8e-93;
Matches 165; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 27 GAPPRLICDSRYLERYLLAKEAENITTCACBSCSNENITVPDTRVNFYAKRMVEVROQ 86
DB 232 GAPPRLICDSRYLORYLLAKEAENITTCACBSCSNENITVPDTRVNFYAKRMVEVROQ 291
OY 87 AVEWOGALLSEAVLRGALLVNSSQPMPEPLOLHYDKAVSGRLSTLTLLRALGAQKEAI 146
DB 292 AVEWOGALLSEAVLRGALLVNSSQPMPEPLOLHYDKAVSGRLSTLTLLRALGAQKEAI 351
OY 147 SPDDASAAPLRTITADTFKRLFRVYSNPLRGKLKLYTGEACRTGDR 193
DB 352 SPDDASAAPLRTITADTFKRLFRVYSNPLRGKLKLYTGEACRTGDR 398

RESULT 4
US-08-318-193-70
; Sequence 70, Application US/08318193
; Patent No. 5641663
; GENERAL INFORMATION:
; APPLICANT: GARYIN, Robert T.
; APPLICANT: MALEK, Lawrence T.
; TITLE OF INVENTION: AN EXPRESSION SYSTEM FOR THE SECRETION
; OF BIOACTIVE HUMAN GRANULOCYTE MACROPHAGE COLONY
; STIMULATING FACTOR (GM-CSF) AND OTHER HETEROLOGOUS
; TITLE OF INVENTION: PROTEINS FROM STREPTOMYCES
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,193
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,314
; FILING DATE:
; APPLICATION NUMBER: US 07/224,568
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 18740/116 CACO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 166 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-318-193-70
Query Match 84.5%; Score 843; DB 1; Length 166;
Best Local Similarity 99.4%; Pred. No. 1.5e-93;
Matches 165; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 28 APPRLCDRYLERLYLEAKAEENITTCGAHCISLNNITVPDTRKVFYANKRMEVROQA 87
Db 1 APPRLCDRYLERLYLEAKAEENITTCGAHCISLNNITVPDTRKVFYANKRMEVROQA 60

QY 88 VEWOGIALLSEAVLRGQALLVNSQWPEPLQLHVDKAVSGLSRLTTLRLALGAKKEAIS 147
Db 61 VEWOGIALLSEAVLRGQALLVNSQWPEPLQLHVDKAVSGLSRLTTLRLALGAKKEAIS 120

QY 148 PPDAASAPLRTITADTFRKLFYVSNFLRGKLLKLTGACRTGDR 193
Db 121 PPDAASAPLRTITADTFRKLFYVSNFLRGKLLKLTGACRTGDR 166

RESULT 5
US-09-604-871-2
; Sequence 2, Application US/09604871
; Patent No. 6340742
; GENERAL INFORMATION:
; APPLICANT: Burg, Josef
; APPLICANT: Hilger, Bernd

APPLICANT: Josef, Hans-Peter
; TITLE OF INVENTION: ERYTHROPOIETIN CONJUGATES
; FILE REFERENCE: 1098 nonprovisional
; CURRENT APPLICATION NUMBER: US/09/604,871
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/151,454
; PRIOR FILING DATE: 1999-08-30
; PRIOR APPLICATION NUMBER: 60/147,452
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: 60/142,243
; PRIOR FILING DATE: 1999-07-02
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-604-871-2
Query Match 84.5%; Score 843; DB 4; Length 166;
Best Local Similarity 99.4%; Pred. No. 1.5e-93;
Matches 165; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 28 APPRLCDRYLERLYLEAKAEENITTCGAHCISLNNITVPDTRKVFYANKRMEVROQA 87
Db 1 APPRLCDRYLERLYLEAKAEENITTCGAHCISLNNITVPDTRKVFYANKRMEVROQA 60

QY 88 VEWOGIALLSEAVLRGQALLVNSQWPEPLQLHVDKAVSGLSRLTTLRLALGAKKEAIS 147
Db 61 VEWOGIALLSEAVLRGQALLVNSQWPEPLQLHVDKAVSGLSRLTTLRLALGAKKEAIS 120

QY 148 PPDAASAPLRTITADTFRKLFYVSNFLRGKLLKLTGACRTGDR 193
Db 121 PPDAASAPLRTITADTFRKLFYVSNFLRGKLLKLTGACRTGDR 166

RESULT 6
PCT-US94-04361-37
; Sequence 37, Application PC/TUS9404361
; GENERAL INFORMATION:
; APPLICANT: Brigham and Women's Hospital
; APPLICANT: 75 Francis Street
; APPLICANT: Boston, MA 02115
; APPLICANT: Bun, H. Franklin
; APPLICANT: Wen, Danyi
; APPLICANT: Showers, Mark O.
; TITLE OF INVENTION: Erythropoietin Muteins With Enhanced
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04361
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/049,802
; FILING DATE: 21-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbal, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 0627.336PC01
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acids
TYPE: amino acid
TOPOLOGY: both
PCT-US94-04361-37

Query Match 84.5%; Score 843; DB 5; Length 166;
Best Local Similarity 99.4%; Pred. No. 1.5e-93;
Matches 165; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 28 APPRLICDSRVLYERLLEKKEAENITTCGAHCISLNEITVPDTKYNFYAMKMEVROQA 87
DB 1 APPRLICDSRVLYERLLEKKEAENITTCGAHCISLNEITVPDTKYNFYAMKMEVROQA 60
QY 88 VEWMGIALISEAVLRGQALLVNSSQWPEPDLQHDKAVSGLSLTLRALGAOKEAIS 147
DB 61 VEWMGIALISEAVLRGQALLVNSSQWPEPDLQHDKAVSGLSLTLRALGAOKEAIS 120
QY 148 PPDASAPLRTITADTFKRLFRVYSNPLRGKIKLYTGACRTGDR 193
DB 121 PPDASAPLRTITADTFKRLFRVYSNPLRGKIKLYTGACRTGDR 166

RESULT 7
US-09-604-871-1
Sequence 1, Application US/09604871
Patent No. 6340742
GENERAL INFORMATION:
APPLICANT: Burig, Josef
APPLICANT: Hilber, Bernd
APPLICANT: Joesel, Hans-Peter
TITLE OF INVENTION: ERYTHROPOIETIN CONJUGATES
FILE REFERENCE: 1098 nonprovisional
CURRENT APPLICATION NUMBER: US/09/604,871
CURRENT FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/151,454
PRIOR FILING DATE: 1999-08-30
PRIOR APPLICATION NUMBER: 60/147,452
PRIOR FILING DATE: 1999-08-05
PRIOR APPLICATION NUMBER: 60/142,243
PRIOR FILING DATE: 1999-07-02
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 165
TYPE: PRT
ORGANISM: Homo sapiens
US-09-604-871-1

Query Match 84.0%; Score 838; DB 4; Length 165;
Best Local Similarity 99.4%; Pred. No. 6.1e-93;
Matches 164; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 28 APPRLICDSRVLYERLLEKKEAENITTCGAHCISLNEITVPDTKYNFYAMKMEVROQA 87
DB 1 APPRLICDSRVLYERLLEKKEAENITTCGAHCISLNEITVPDTKYNFYAMKMEVROQA 60
QY 88 VEWMGIALISEAVLRGQALLVNSSQWPEPDLQHDKAVSGLSLTLRALGAOKEAIS 147
DB 61 VEWMGIALISEAVLRGQALLVNSSQWPEPDLQHDKAVSGLSLTLRALGAOKEAIS 120
QY 148 PPDASAPLRTITADTFKRLFRVYSNPLRGKIKLYTGACRTGDR 192
DB 121 PPDASAPLRTITADTFKRLFRVYSNPLRGKIKLYTGACRTGDR 165

RESULT 8
PCT-US94-04361-45
Sequence 45, Application PC/TUS9404361
GENERAL INFORMATION:

APPLICANT: Brigham and Women's Hospital
APPLICANT: 75 Francis Street
APPLICANT: Boston, MA 02115
APPLICANT: Bunn, H. Franklin
APPLICANT: Men, Danyl
TITLE OF INVENTION: Erythropoietin Mutelins With Enhanced
TITLE OF INVENTION: Activity
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04361
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/049,802
FILING DATE: 21-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Cimbala, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0627.336PC01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acids
TYPE: amino acid
TOPOLOGY: both
PCT-US94-04361-45

Query Match 83.7%; Score 835; DB 5; Length 166;
Best Local Similarity 98.8%; Pred. No. 1.4e-92;
Matches 164; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 28 APPRLICDSRVLYERLLEKKEAENITTCGAHCISLNEITVPDTKYNFYAMKMEVROQA 87
DB 1 APPRLICDSRVLYERLLEKKEAENITTCGAHCISLNEITVPDTKYNFYAMKMEVROQA 60
QY 88 VEWMGIALISEAVLRGQALLVNSSQWPEPDLQHDKAVSGLSLTLRALGAOKEAIS 147
DB 61 VEWMGIALISEAVLRGQALLVNSSQWPEPDLQHDKAVSGLSLTLRALGAOKEAIS 120
QY 148 PPDASAPLRTITADTFKRLFRVYSNPLRGKIKLYTGACRTGDR 193
DB 121 PPDASAPLRTITADTFKRLFRVYSNPLRGKIKLYTGACRTGDR 166

RESULT 9
PCT-US94-04361-38
Sequence 38, Application PC/TUS9404361
GENERAL INFORMATION:
APPLICANT: Brigham and Women's Hospital
APPLICANT: 75 Francis Street
APPLICANT: Boston, MA 02115
APPLICANT: Bunn, H. Franklin
APPLICANT: Men, Danyl
TITLE OF INVENTION: Erythropoietin Mutelins With Enhanced
TITLE OF INVENTION: Activity
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:

ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04361
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/049,802
FILING DATE: 21-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Cimbala, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0627.336PC01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 165 amino acids
TYPE: amino acid
TOPOLOGY: both
PCT-US94-04361-38

Query Match 76.3%; Score 761.5; DB 5; Length 165;
Best Local Similarity 91.0%; Pred. No. 1e-83;
Matches 151; Conservative 7; Mismatches 7; Indels 1; Gaps 1;

QY 28 APRRLCDSEVLERLYLLEAKEAENITTCGAHCSLNENITVPDTKVFYAMKREVEVQQA 87
DB 1 APRRLCDSEVLERLYLLEAKEAENITTCGAHCSLNENITVPDTKVFYAMKREVEVQQA 60

QY 88 VEWQGLALISEAVLRQALLVNSQPEWPELOLHVDAVSGLSLTLRALGAKREAFIS 147
DB 61 VEWQGLALISEAVLRQALLVNSQPEWPELOLHVDAVSGLSLTLRALGAKREAFIS 119

QY 148 PPDASAAPLRTITADTFKRLFRVYSNPLRGKLTLYGECRCRGR 193
DB 120 LPDASAAPLRTITADTFKRLFRVYSNPLRGKLTLYGECRCRGR 165

RESULT 10
PCT-US94-04361-39
Sequence 39, Application PC/TUS9404361
GENERAL INFORMATION:
APPLICANT: Brigham and Women's Hospital
APPLICANT: 75 Francis Street
APPLICANT: Boston, MA 02115
APPLICANT: Bunn, H. Franklin
APPLICANT: Wen, Danyl
APPLICANT: Showers, Mark O.
TITLE OF INVENTION: Erythropoietin Mutelins With Enhanced
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04361
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/049,802
FILING DATE: 21-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Cimbala, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0627.336PC01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 165 amino acids
TYPE: amino acid
TOPOLOGY: both
PCT-US94-04361-39

Query Match 75.8%; Score 756.5; DB 5; Length 165;
Best Local Similarity 89.8%; Pred. No. 4e-83;
Matches 149; Conservative 9; Mismatches 7; Indels 1; Gaps 1;

QY 28 APRRLCDSEVLERLYLLEAKEAENITTCGAHCSLNENITVPDTKVFYAMKREVEVQQA 87
DB 1 APRRLCDSEVLERLYLLEAKEAENITTCGAHCSLNENITVPDTKVFYAMKREVEVQQA 60

QY 88 VEWQGLALISEAVLRQALLVNSQPEWPELOLHVDAVSGLSLTLRALGAKREAFIS 147
DB 61 VEWQGLALISEAVLRQALLVNSQPEWPELOLHVDAVSGLSLTLRALGAKREAFIS 119

QY 148 PPDASAAPLRTITADTFKRLFRVYSNPLRGKLTLYGECRCRGR 193
DB 120 LPDASAAPLRTITADTFKRLFRVYSNPLRGKLTLYGECRCRGR 165

RESULT 11
PCT-US94-04361-44
Sequence 44, Application PC/TUS9404361
GENERAL INFORMATION:
APPLICANT: Brigham and Women's Hospital
APPLICANT: 75 Francis Street
APPLICANT: Boston, MA 02115
APPLICANT: Bunn, H. Franklin
APPLICANT: Wen, Danyl
APPLICANT: Showers, Mark O.
TITLE OF INVENTION: Erythropoietin Mutelins With Enhanced
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04361
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/049,802
FILING DATE: 21-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Cimbala, Michele A.

```

? REGISTRATION NUMBER: 33,851
? REFERENCE/DOCKET NUMBER: 0627.336PC01
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (202) 371-2600
? TELEFAX: (202) 371-2540
? INFORMATION FOR SEQ ID NO: 44:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 166 amino acids
? TYPE: amino acid
? TOPOLOGY: both
PCT-US94-04361-44

Query Match          71.1%; Score 710; DB 5; Length 166;
Best Local Similarity 83.7%; Pred. No.1,6e-77;
Matches 139; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

Oy 28 APRRLICDSRVLEERYLLLEAKAEANTTGCAGHCSTINENTVPTDKVFYAMRMEYRQQA 87
Db 1 APRRLICDSRVLEERYLLLEAREANVTMGCAEGCSSENITVPDTKNFYWMRDVGQA 60
Oy 88 VEVWGGLALISAVLRGQALLVNSQPWEPDLQHDYKAVSGLSLTTLRALGAOKRAIS 147
Db 61 VEVWGGLALISAVLRGQALLVNSQSPEETQLDHYDAVSSLSRLSLRALGAOKRAIS 120
Oy 148 PDASAAPIRTFTADTFPKLFRVYSNPLRGKLXTGTGEACHTGDR 193
Db 121 LPEATSAPLRTFTVDYLCKLPRIITSNFLRGKLFLLTGACRRGRD 166

RESULT 12
PCT-US94-04361-41
Sequence 41, Application PC/TUS9404361
GENERAL INFORMATION:
APPLICANT: Brigham and Women's Hospital
APPLICANT: 75 Francis Street
APPLICANT: Boston, MA 02115
APPLICANT: Bunn, H. Franklin
APPLICANT: Men, Danyl
APPLICANT: Showers, Mark O.
TITLE OF INVENTION: Erythropoietin Mutelins With Enhanced
TITLE OF INVENTION: Activity
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04361
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/049,802
FILING DATE: 21-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Gimbalda, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0627.336PC01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 41:
LENGTH: 166 amino acids
TYPE: amino acid
TOPOLOGY: both

```

```

PCT-US94-04361-41
Query Match      70.9%; Score 708; DB 5; Length 166;
Best Local Similarity 82.5%; Pred. No. 2.8e-77;
Matches 137; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

QY 28 APPRICDSRVLERLTLEKAEANTTGCAEHCISINENITVPDTKVNFAMRMEVROOA 87
   |||||.....:|||..|||.:|||||.....:|
DB 1 APPRICDSRVLERLTLEKAEANTMGCAEGEPRESNITTVPDTKVNFAMRMEVEEOA 60
   |||||.....:|||..|||.:|||||.....:|
QY 88 VEWOGLLALSAVLRGALLVNSSQPWEPIQLHVDKAVSGLRSITLTLRALGAQKEAIS 147
   |||||.....:|||..|||.:|||||.....:|
DB 61 VEWMGSLLSLSEAILLOAQMLCANSSQPEPSLIOLHIDKAISRISITSLRYLAGAKELMS 120
   |||||.....:|||..|||.:|||||.....:|
QY 148 PPDASAAAPLRITTTADTFPRKLFRVYSNFRGLKLYTGACTGTGR 193
   |||||.....:|||..|||.:|||||.....:|
DB 121 PPDATQAAPLRILTADTFCKLFRRVYSNFLRGLKLYTGACRRGRD 166
   |||||.....:|||..|||.:|||||.....:|

RESULT 13
PCT-US94-04361-40
; Sequence 40, Application PC/TUS9404361
; GENERAL INFORMATION:
; APPLICANT: Brigham and Women's Hospital
; APPLICANT: 75 Francis Street
; APPLICANT: Boston, MA 02115
; APPLICANT: Bunn, H. Franklin
; APPLICANT: Men, Danyl
; APPLICANT: Showers, Mark O.
; TITLE OF INVENTION: Erythropoietin Mutelins With Enhanced
; TITLE OF INVENTION: Activity
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04361
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/049,802
; FILING DATE: 21-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbalia, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 0627.336PC01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 166 amino acids
; TYPE: amino acid
; TOPOLOGY: both
; PCT-US94-04361-40

Query Match      69.7%; Score 696; DB 5; Length 166;
Best Local Similarity 80.1%; Pred. No. 7.8e-76;
Matches 133; Conservative 14; Mismatches 19; Indels 0; Gaps 0;

QY 28 APPRICDSRVLERLTLEKAEANTTGCAEHCISINENITVPDTKVNFAMRMEVROOA 87
   |||||.....:|||..|||.:|||||.....:|
DB 1 APPRICDSRVLERLTLEKAEANTMGCAEGEPRESNITTVPDTKVNFAMRMEVEEOA 60
   |||||.....:|||..|||.:|||||.....:|

```

[illegible]

RESULT 14

```

1      PCT-US94-04361-42
2      Sequence 42, Application PC/US94/04361
3      GENERAL INFORMATION:
4      APPLICANT: Brigham and Women's Hospital
5      APPLICANT: 75 Francis Street
6      APPLICANT: Boston, MA 02115
7      APPLICANT: Bunn, H. Franklin
8      APPLICANT: Wen, Danyl
9      APPLICANT: Showers, Mark O.
10     TITLE OF INVENTION: Erythropoietin Mutelins With Enhanced
11     TITLE OF INVENTION: Activity
12     NUMBER OF SEQUENCES: 59
13     CORRESPONDENCE ADDRESSES:
14     ADDRESSEE: Sterne, Kessler, Goldstein & Fox
15     STREET: 1100 New York Avenue, Suite 600
16     CITY: Washington
17     STATE: D.C.
18     COUNTRY: U.S.A.
19     ZIP: 20005-3934
20     COMPUTER READABLE FORM:
21     MEDIUM TYPE: Floppy disk
22     COMPUTER: IBM PC Compatible
23     OPERATING SYSTEM: PC-DOS/MS-DOS
24     SOFTWARE: PatentIn Release #1.0, Version #1.25
25     CURRENT APPLICATION DATA:
26     APPLICATION NUMBER: PCT/US94/04361
27     FILING DATE: Herewith
28     CLASSIFICATION:
29     PRIOR APPLICATION DATA:
30     APPLICATION NUMBER: 08/049,802
31     FILING DATE: 21-APR-1993
32     ATTORNEY/AGENT INFORMATION:
33     NAME: Cimbala, Michele A.
34     REGISTRATION NUMBER: 33,851
35     REFERENCE/DOCKET NUMBER: 0627,336PC01
36     TELECOMMUNICATION INFORMATION:
37     TELEPHONE: (202) 371-2600
38     TELEFAX: (202) 371-2540
39     INFORMATION FOR SEQ ID NO: 42:
40     SEQUENCE CHARACTERISTICS:
41     LENGTH: 167 amino acids
42     TYPE: amino acid
43     TOPOLOGY: both
44     PCT-US94-04361-42

```

Query Match	68.8%;	Score 686.5;	DB 5;	Length 167;
Best Local Similarity	81.4%;	Pred. No. 11e-74;		
Matches 136;	Conservative 10;	Mismatches 20;	Indels 1;	Gaps 1
QY	28	APPRICDSRYLERYLTAKEAEANTTTGCAEHCISINENITVDTKYNFAMKRMVEYRQOA	87	
Db	1	APPRICDSRYLERIILTAKEAENNTMGCAECCSPSENITVDTKYNFAMKRMVEYQOA	60	
QY	88	VEWOGALLSEAVLRGQALLVNSQPMWEPPLQAHWDKAVSGLSRTLLTALGAQKEAITS	147	
Db	61	LEWOGALLSEALIFRGQALPANASQPCFALRLHWDKAVSGLSRTSLTALGAQKEALP	120	
QY	148	PPDAA-SAAPLRITITADTFRKLFRRYYSNLRKRLKLYLGEACRTGDR	193	
Db	121	LPDAPSAAPLRITFVDALSKLERTYSNLRKRLKLYLGEACRRGR	167	

RESULT 15

```

PCT-US94-04361-43
Sequence 43, Application PC/TUS9404361
GENERAL INFORMATION:
APPLICANT: Bingham and Women's Hospital
APPLICANT: 75 Francis Street
APPLICANT: Boston, MA 02115
APPLICANT: Bunn, H. Franklin
APPLICANT: Wen, Danyl
APPLICANT: Showers, Mark O.
TITLE OF INVENTION: Erythropoietin Muteins With Enhanced
TITLE OF INVENTION: Activity
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04361
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/049,802
FILING DATE: 21-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Cimbal, Michele A.
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0627.336PC01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 168 amino acids
TYPE: amino acid
TOPOLOGY: both
PCT-US94-04361-43

```

Query Match 68.7%; Score 686; DB 5; Length 168;
Best Local Similarity 82.1%; Pred. No. 1.3e-74;
Matches 138; Conservative 8; Mismatches 20; Indels 2; Gaps 1

QY 28 APPRLICDSRYLERYLLAKEAEANITTTGCAEHCSTINENITVDTKYNFPAAMRMEYQQQA 87
Db 1 APPRICDSRYLERITLEKEGENNTMGACSCSPSENTIVPTDKNFPAMRMEYQQQA 60

QY 88 VEEWOGIALLSAVLRGQALLVNSSQPWFPIOLHYDRAVSGIRSLTTLRALCAQKEAITS 147
Db 61 MEWVGIALLSAIIQGQALLANNSSQPSALIOHVDKAASGIRSLTSLRALGAQKEAIRP 120

QY 148 PPDA--ASAPLRTTADTFERKLFRVYSNFELRGKLTLYGEACRGSDR 193
Db 121 LPDASPSSATPLRTVAVDTLCKLFNNYSFLNGKLTLYTGECRRDR 168

Search completed: January 8, 2003, 05:16:25
Job time : 18.9025 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 03:41:38 ; Search time 33.3315 Seconds
(without alignments)
112.388 Million cell updates/sec

Title: US-09-813-775c-2

Perfect score: 998

Sequence: 1 MCVHECPAMWLILSLSLP.....NFLRGKLYTGECRTGDR 193

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 118974 seqs, 19401057 residues

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*

- 1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84.3	84.5	166	10	US-09-853-731-2
2	84.3	84.5	166	12	US-10-014-363-2
3	84.3	84.5	169	12	US-10-014-363-4
4	84.3	84.5	174	12	US-10-014-363-3
5	84.3	84.5	174	12	US-10-014-363-5
6	83.8	84.0	165	10	US-09-853-731-1
7	83.8	84.0	165	12	US-10-014-363-1
8	28.6	28.7	67	10	US-09-864-761-48502
9	14.8	14.8	27	10	US-09-864-761-404-11
10	11.1	11.1	30	10	US-09-975-063-2
11	11.1	11.1	30	12	US-10-011-858-2
12	10.8	10.8	20	9	US-09-919-703-9
13	8.8	8.8	20	9	US-09-919-703-10
14	8.1	8.1	147	10	US-09-729-674-279
15	77.5	7.8	813	9	US-09-964-899-25
16	7.6	7.6	1005	9	US-09-988-117-1
17	7.6	7.6	1005	10	US-09-812-471-1
18	7.4	7.4	1005	10	US-09-812-633-1
19	7.4	7.4	136	10	US-09-764-877-1623

20	7.4	7.4	2201	12	US-10-029-907-3	Sequence 3, Appl1
21	73.5	7.4	210	10	US-09-731-872-387	Sequence 387, App
22	73.5	7.4	418	10	US-09-992-964-14	Sequence 14, Appl
23	73.5	7.4	468	10	US-09-887-879-14	Sequence 14, Appl
24	73.5	7.4	468	12	US-10-039-785-1	Sequence 11, Appl
25	7.2	7.2	224	10	US-09-242-980-11	Sequence 11, Appl
26	7.2	7.2	824	9	US-09-932-145-3	Sequence 3, Appl1
27	7.2	7.2	81	10	US-09-764-877-1947	Sequence 1947, Ap
28	71.5	7.2	311	10	US-09-191-687B-5	Sequence 5, Appl1
29	71.5	7.2	1643	9	US-09-738-626-6894	Sequence 6894, Ap
30	7.1	7.1	283	10	US-09-972-599A-55	Sequence 55, Appl
31	7.1	7.1	473	9	US-09-978-295A-400	Sequence 400, App
32	7.1	7.1	473	9	US-09-978-697-400	Sequence 400, App
33	7.1	7.1	473	9	US-09-978-192A-400	Sequence 400, App
34	7.1	7.1	473	9	US-09-999-832A-400	Sequence 400, App
35	7.1	7.1	473	9	US-09-978-189-400	Sequence 400, App
36	7.1	7.1	473	9	US-10-028-072-382	Sequence 382, App
37	7.1	7.1	473	10	US-09-758-140-2	Sequence 2, Appl1
38	7.1	7.1	473	10	US-09-796-858-34	Sequence 34, Appl
39	7.1	7.1	473	10	US-09-893-348-26	Sequence 26, Appl
40	7.1	7.1	473	10	US-09-972-599A-2	Sequence 2, Appl1
41	70.5	7.1	614	10	US-09-803-589-4	Sequence 4, Appl1
42	7.0	7.0	362	10	US-09-731-872-280	Sequence 280, App
43	7.0	7.0	489	9	US-09-992-598-138	Sequence 138, App
44	7.0	7.0	489	9	US-09-989-293A-138	Sequence 138, App
45	7.0	7.0	489	9	US-09-989-735-138	Sequence 138, App

ALIGNMENTS

RESULT 1
US-09-853-731-2
Sequence 2, Application US/09853731
Patient No. US20020037841A1
GENERAL INFORMATION:
APPLICANT: Papadimitriou, Apollon
TITLE OR INVENTION: Erythropoietin Composition
FILE REFERENCE: 20619 US
CURRENT APPLICATION NUMBER: US/09/853,731
CURRENT FILING DATE: 2001-05-11
PRIOR APPLICATION NUMBER: EP/00110355.5
PRIOR FILING DATE: 2000-05-15
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 166
TYPE: PRT
ORGANISM: Homo sapiens
US-09-853-731-2

Query Match 84.5%; Score 843; DB 10; Length 166;
Best Local Similarity 99.4%; Pred. No. 1.3e-79;
Matches 165; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 28 APPPLICDSRYLEBYLLEAKAEENITTCAGHCSLNENITVPDTKVNFPYAKRMEVROQA 87
DB 1 APPPLICDSRYLEBYLLEAKAEENITTCAGHCSLNENITVPDTKVNFPYAKRMEVROQA 60
QY 88 VEVWOGIALISEAVLRQALIVNSOPWEPIQLHVDKAVSGLSLTTLRALGAQKEAIS 147
DB 61 VEVWOGIALISEAVLRQALIVNSOPWEPIQLHVDKAVSGLSLTTLRALGAQKEAIS 120
QY 148 PPDASAAPLRTITADPFRKLEFRYNSPLRGKLYTGECRTGDR 193
DB 121 PPDASAAPLRTITADPFRKLEFRYNSPLRGKLYTGECRTGDR 166

RESULT 2
US-10-014-363-2
Sequence 2, Application US/10014363
Patient No. US20020115833A1
GENERAL INFORMATION:

APPLICANT: Burg, Josef
APPLICANT: Engel, Alfred
APPLICANT: Franze, Reinhard
APPLICANT: Hilger, Bernd
APPLICANT: Schurig, Hartmut Ernst
APPLICANT: Tischer, Wilhelm
APPLICANT: Wozny, Manfred
TITLE OF INVENTION: Erythropoietin Conjugates
FILE REFERENCE: Case 20805
CURRENT APPLICATION NUMBER: US/10/014,363
CURRENT FILING DATE: 2001-12-11
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 166
TYPE: PRT
ORGANISM: Homo sapiens
US-10-014-363-2

Query Match 84.5%; Score 843; DB 12; Length 166;
Best Local Similarity 99.4%; Pred. No. 1.3e-79;
Matches 165; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

28 APRRLICDSRVLEERYLLEAKENITTCGAHCISLNIENITVPDTKVFYAMKMEYRQQA 87
1 APRRLICDSRVLEERYLLEAKENITTCGAHCISLNIENITVPDTKVFYAMKMEYRQQA 60
88 VEWOGALLSEAVLRGQALLVNSSQPWEPLQHLVDKAVSGRLSTLTTLRALGAOKREAIS 147
61 VEWOGALLSEAVLRGQALLVNSSQPWEPLQHLVDKAVSGRLSTLTTLRALGAOKREAIS 120
148 PPDASAPLRTITADTFKRLFRVYSNPLRGKILKLTGTGACRTGDR 193
121 PPDASAPLRTITADTFKRLFRVYSNPLRGKILKLTGTGACRTGDR 166

RESULT 3

US-10-014-363-4
Sequence 4, Application US/10014363
Patent No. US20020115833A1
GENERAL INFORMATION:
APPLICANT: Burg, Josef
APPLICANT: Engel, Alfred
APPLICANT: Franze, Reinhard
APPLICANT: Hilger, Bernd
APPLICANT: Schurig, Hartmut Ernst
APPLICANT: Tischer, Wilhelm
APPLICANT: Wozny, Manfred
TITLE OF INVENTION: Erythropoietin Conjugates
FILE REFERENCE: Case 20805
CURRENT APPLICATION NUMBER: US/10/014,363
CURRENT FILING DATE: 2001-12-11
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 169
TYPE: PRT
ORGANISM: CHO/dhfr-
US-10-014-363-4

Query Match 84.5%; Score 843; DB 12; Length 169;
Best Local Similarity 99.4%; Pred. No. 1.4e-79;
Matches 165; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

28 APRRLICDSRVLEERYLLEAKENITTCGAHCISLNIENITVPDTKVFYAMKMEYRQQA 87
4 APRRLICDSRVLEERYLLEAKENITTCGAHCISLNIENITVPDTKVFYAMKMEYRQQA 63
88 VEWOGALLSEAVLRGQALLVNSSQPWEPLQHLVDKAVSGRLSTLTTLRALGAOKREAIS 147
64 VEWOGALLSEAVLRGQALLVNSSQPWEPLQHLVDKAVSGRLSTLTTLRALGAOKREAIS 123
148 PPDASAPLRTITADTFKRLFRVYSNPLRGKILKLTGTGACRTGDR 193

124 PPDASAPLRTITADTFKRLFRVYSNPLRGKILKLTGTGACRTGDR 169

RESULT 4
US-10-014-363-3
Sequence 3, Application US/10014363
Patent No. US20020115833A1
GENERAL INFORMATION:
APPLICANT: Burg, Josef
APPLICANT: Engel, Alfred
APPLICANT: Franze, Reinhard
APPLICANT: Hilger, Bernd
APPLICANT: Schurig, Hartmut Ernst
APPLICANT: Tischer, Wilhelm
APPLICANT: Wozny, Manfred
TITLE OF INVENTION: Erythropoietin Conjugates
FILE REFERENCE: Case 20805
CURRENT APPLICATION NUMBER: US/10/014,363
CURRENT FILING DATE: 2001-12-11
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 174
TYPE: PRT
ORGANISM: CHO/dhfr-
US-10-014-363-3

Query Match 84.5%; Score 843; DB 12; Length 174;
Best Local Similarity 99.4%; Pred. No. 1.4e-79;
Matches 165; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

28 APRRLICDSRVLEERYLLEAKENITTCGAHCISLNIENITVPDTKVFYAMKMEYRQQA 87
9 APRRLICDSRVLEERYLLEAKENITTCGAHCISLNIENITVPDTKVFYAMKMEYRQQA 68
88 VEWOGALLSEAVLRGQALLVNSSQPWEPLQHLVDKAVSGRLSTLTTLRALGAOKREAIS 147
69 VEWOGALLSEAVLRGQALLVNSSQPWEPLQHLVDKAVSGRLSTLTTLRALGAOKREAIS 128
148 PPDASAPLRTITADTFKRLFRVYSNPLRGKILKLTGTGACRTGDR 193
129 PPDASAPLRTITADTFKRLFRVYSNPLRGKILKLTGTGACRTGDR 174

RESULT 5

US-10-014-363-5
Sequence 5, Application US/10014363
Patent No. US20020115833A1
GENERAL INFORMATION:
APPLICANT: Burg, Josef
APPLICANT: Engel, Alfred
APPLICANT: Franze, Reinhard
APPLICANT: Hilger, Bernd
APPLICANT: Schurig, Hartmut Ernst
APPLICANT: Tischer, Wilhelm
APPLICANT: Wozny, Manfred
TITLE OF INVENTION: Erythropoietin Conjugates
FILE REFERENCE: Case 20805
CURRENT APPLICATION NUMBER: US/10/014,363
CURRENT FILING DATE: 2001-12-11
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 174
TYPE: PRT
ORGANISM: CHO/dhfr-
US-10-014-363-5

Query Match 84.5%; Score 843; DB 12; Length 174;
Best Local Similarity 99.4%; Pred. No. 1.4e-79;
Matches 165; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 28 APRRLICSRVLERLLEAKAEANITTCGAHCISLNETITVPDTKVFAMKRMVEYROOA 87
Db 9 APRRLICSRVLERLLEAKAEANITTCGAHCISLNETITVPDTKVFAMKRMVEYGOOA 68
OY 88 VEWMOGLALLSEAVLRGQALLVNSSQPWEPLDLHVDKAVSGRLSITTLRLALGAOKEATS 147
Db 69 VEWMOGLALLSEAVLRGQALLVNSSQPWEPLDLHVDKAVSGRLSITTLRLALGAOKEATS 128
OY 148 PPDAASAPLRTITADTFRKLFRRVYSNPLRGKLTLYTGACRTGDR 193
Db 129 PPDAASAPLRTITADTFRKLFRRVYSNPLRGKLTLYTGACRTGDR 174

RESULT 6
US-09-853-731-1
: Sequence 1, Application US/09853731
: Patent No. US20020037841A1
: GENERAL INFORMATION:
: APPLICANT: Papadimitriou, Apollon
: TITLE OF INVENTION: Erythropoietin Composition
: FILE REFERENCE: 20619 US
: CURRENT APPLICATION NUMBER: US/09/853,731
: PRIOR APPLICATION NUMBER: EP/00110355.5
: PRIOR FILING DATE: 2000-05-15
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 1
: LENGTH: 165
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-853-731-1

Query Match 84.0%; Score 838; DB 10; Length 165;
Best Local Similarity 99.4%; Pred. No. 4,3e-79;
Matches 164; Conservative 0; Mismatches 1; Indels 0; Gaps 0

OY 28 APRRLICSRVLERLLEAKAEANITTCGAHCISLNETITVPDTKVFAMKRMVEYROOA 87
Db 1 APRRLICSRVLERLLEAKAEANITTCGAHCISLNETITVPDTKVFAMKRMVEYGOOA 60
OY 88 VEWMOGLALLSEAVLRGQALLVNSSQPWEPLDLHVDKAVSGRLSITTLRLALGAOKEATS 147
Db 61 VEWMOGLALLSEAVLRGQALLVNSSQPWEPLDLHVDKAVSGRLSITTLRLALGAOKEATS 120
OY 148 PPDAASAPLRTITADTFRKLFRRVYSNPLRGKLTLYTGACRTGD 192
Db 121 PPDAASAPLRTITADTFRKLFRRVYSNPLRGKLTLYTGACRTGD 165

RESULT 7
US-10-014-363-1
: Sequence 1, Application US/10014363
: Patent No. US20020115833A1
: GENERAL INFORMATION:
: APPLICANT: Burg, Josef
: APPLICANT: Engel, Alfred
: APPLICANT: Franze, Reinhard
: APPLICANT: Hilger, Bernd
: APPLICANT: Schurig, Hartmut Ernst
: APPLICANT: Fischer, Wilhelm
: APPLICANT: Moznay, Manfred
: TITLE OF INVENTION: Erythropoietin Conjugates
: FILE REFERENCE: Case 20805
: CURRENT APPLICATION NUMBER: US/10/014,363
: CURRENT FILING DATE: 2001-12-11
: NUMBER OF SEQ ID NOS: 5
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1
: LENGTH: 165
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-014-363-1

```

```

Query Match          84.0%: Score 838; DB 12; Length 165;
Best Local Similarity 99.4%: Pred. No. 4.3e-79;
Matches 164; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      28  APRRLICDSRYLERYLLERLLAKEAEINITTGCAEHCISINENITVPDTRKYNFYANKRMEVROQA 87
DB      1  APRRLICDSRYLERLLERLLAKEAEINITTGCAEHCISINENITVPDTRKYNFYANKRMEVGOQA 60

QY      88  VEVWOGALLISEAVLRGALLVNSSQPEWPIQLHYDKAVSGLRISITTLRLACAKQKKAIS 147
DB      61  VEVWOGALLISEAVLRGALLVNSSQPEWPIQLHYDKAVSGLRISITTLRLALGAQKKAIS 120

QY      148  PPDAASAPLRTITPDTRKLFERYYSNPLRGKRLKLTGECRTGD 192
DB      121  PPDAASAPLRTITPDTRKLFERYYSNPLRGKRLKLTGECRTGD 165

RESULT 8
US-09-864-761-48502
: Sequence 48502, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
:   APPLICANT: Penn, Sharon G.
:   APPLICANT: Rank, David R.
:   APPLICANT: Hanzel, David K.
:   APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEOTIC ACID PROBES USEFUL FOR
: FILE REFERENCE: Aecolica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
:   CURRENT FILING DATE: 2001-05-23
:   PRIOR APPLICATION NUMBER: US 60/180,312
:   PRIOR FILING DATE: 2000-02-04
:   PRIOR APPLICATION NUMBER: US 60/207,456
:   PRIOR FILING DATE: 2000-05-26
:   PRIOR APPLICATION NUMBER: US 09/632,366
:   PRIOR FILING DATE: 2000-08-03
:   PRIOR APPLICATION NUMBER: GB 24263.6
:   PRIOR FILING DATE: 2000-10-04
:   PRIOR APPLICATION NUMBER: US 60/236,359
:   PRIOR FILING DATE: 2000-09-27
:   PRIOR APPLICATION NUMBER: PCT/US01/00666
:   PRIOR FILING DATE: 2001-01-30
:   PRIOR APPLICATION NUMBER: PCT/US01/00667
:   PRIOR FILING DATE: 2001-01-30
:   PRIOR APPLICATION NUMBER: PCT/US01/00664
:   PRIOR FILING DATE: 2001-01-30
:   PRIOR APPLICATION NUMBER: PCT/US01/00669
:   PRIOR FILING DATE: 2001-01-30
:   PRIOR APPLICATION NUMBER: PCT/US01/00665
:   PRIOR FILING DATE: 2001-01-30
:   PRIOR APPLICATION NUMBER: PCT/US01/00668
:   PRIOR FILING DATE: 2001-01-30
:   PRIOR APPLICATION NUMBER: PCT/US01/00663
:   PRIOR FILING DATE: 2001-01-30
:   PRIOR APPLICATION NUMBER: PCT/US01/00662
:   PRIOR FILING DATE: 2001-01-30
:   PRIOR APPLICATION NUMBER: PCT/US01/00661
:   PRIOR FILING DATE: 2001-01-30
:   PRIOR APPLICATION NUMBER: PCT/US01/00670
:   PRIOR FILING DATE: 2001-01-30
:   PRIOR APPLICATION NUMBER: US 60/234,667
:   PRIOR FILING DATE: 2000-09-21
:   PRIOR APPLICATION NUMBER: US 09/608,408
:   PRIOR FILING DATE: 2000-06-30
:   PRIOR APPLICATION NUMBER: US 09/774,203
:   PRIOR FILING DATE: 2001-01-29
:   NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 48502
: LENGTH: 60
: TYPE: PRT

```

ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AF053356.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.93
OTHER INFORMATION: SWISSPROT HIT: P01588, EVALUE 9.00e-28
OTHER INFORMATION: EST_HUMAN HIT: AA662379.1, EVALUE 3.00e-10
US-09-864-761-48502

Query Match 28.7%; Score 286; DB 10; Length 60;
Best Local Similarity 98.3%; Pred. No. 7.3e-23;
Matches 59; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 83 VROQAEVWQGLALISEAVLRGALLVNSQPEPQLQHYDKAVSGLRSTLTLLRAGAQ 142
DB 1 VGGQAAVEWQGLALISEAVLRGALLVNSQPEPQLQHYDKAVSGLRSTLTLLRAGAQ 60

RESULT 9

US-09-886-404-11
Sequence 11, Application US/09886404
Patent No. US20020037524A1
GENERAL INFORMATION:
APPLICANT: Medlock, Eugene
APPLICANT: Yeh, Richard
APPLICANT: Silbiger, Scott M.
APPLICANT: Elliot, Gary S.
APPLICANT: Nguyen, Hung Q.
TITLE OF INVENTION: IL-17 Like Molecules and Uses Thereof
FILE REFERENCE: 01017/37128B
CURRENT APPLICATION NUMBER: US/09/886,404
CURRENT FILING DATE: 2001-06-21
PRIOR APPLICATION NUMBER: 09/810,384
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/266,159
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/213,125
PRIOR FILING DATE: 2000-06-22
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 11
LENGTH: 27
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Epogen signal
US-09-886-404-11

Query Match 14.8%; Score 148; DB 10; Length 27;
Best Local Similarity 100.0%; Pred. No. 4e-09;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MGVHECPAMWLILSLSLPLGLPVLG 27
DB 1 MGVHECPAMWLILSLSLPLGLPVLG 27

RESULT 10
US-09-975-063-2
Sequence 2, Application US/09975063
Patent No. US20020045255A1
GENERAL INFORMATION:
APPLICANT: POWELL, Jerry S.
TITLE OF INVENTION: HUMAN ERYTHROPOIETIN GENE: HIGH LEVEL
EXPRESSION IN STABLY TRANSFECTED MAMMALIAN CELLS

NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dorsey & Whitney, LLP
STREET: Suite 3400, 1420 Fifth Avenue, U.S. Bank Centre
CITY: Seattle
STATE: Washington
COUNTRY: USA

ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.7/7.8
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/975,063
FILING DATE: 10-Oct-2001
ATTORNEY/AGENT INFORMATION:
NAME: Roberts, Mark W.
REGISTRATION NUMBER: 46,160
REFERENCE/DOCKET NUMBER: 500582.03 (112893.109)

TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-903-8728
TELEFAX: 206-903-8820

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-975-063-2

Query Match 11.1%; Score 111; DB 10; Length 30;
Best Local Similarity 83.3%; Pred. No. 3e-05;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 28 APRRLICDSRYLERYLLEKAEKNTTGA 57
DB 1 APRRLICDSRYLERYLLEKAEKNTTGA 30

RESULT 11

US-10-011-858-2
Sequence 2, Application US/10011858
Patent No. US20020137145A1
GENERAL INFORMATION:
APPLICANT: POWELL, Jerry S.
TITLE OF INVENTION: HUMAN ERYTHROPOIETIN GENE: HIGH LEVEL
EXPRESSION IN STABLY TRANSFECTED MAMMALIAN CELLS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dorsey & Whitney, LLP
STREET: Suite 3400, 1420 Fifth Avenue, U.S. Bank Centre
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS-WORD
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/011,858
FILING DATE: 05-NO. US20020137145A1-2001
ATTORNEY/AGENT INFORMATION:
NAME: Roberts, Mark W.
REGISTRATION NUMBER: 46,160
REFERENCE/DOCKET NUMBER: 500582.13
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-903-8728
TELEFAX: 206-903-8820

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide

```

;          SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-011-858-2

```

```

Query Match      11.1%; Score 111; DB 12; Length 30;
Best Local Similarity 83.3%; Pred. No. 3e-05;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```
QY      28 APPRLICDSRVLERYLLAEKAENITTCGA   57
          |||||
Db       1 APXRLILDSRVLERYLLAEKAEEXITDCCA   30
```

RESULT 12
US-09-919-703-9

```

; APPLICANT: Kyslal, Simon W.
; APPLICANT: Rabin, Gerald
; TITLE OF INVENTION: Peptides and Their Use to Ameliorate
; TITLE OF INVENTION: Cell Death
; FILE REFERENCE: 50216/003004
; CURRENT APPLICATION NUMBER: US/09/919,703
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 09/294,457
; PRIOR FILING DATE: 1999-04-19
; PRIOR APPLICATION NUMBER: US 08/759,599
; PRIOR FILING DATE: 1996-12-05
; PRIOR APPLICATION NUMBER: US 60/008,233
; PRIOR FILING DATE: 1995-12-06
; NUMBER OF SEQ. ID NOS.: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
; US-09-919-703-9

```

```
Query Match      10.8%; Score 108; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      174 NPLRGKLLKYTGACRTGDR 193  
          |||  
Db       1 NFLRGLKLKYTGACRTGDR 20
```

```

RESULT 13
US-09-919-703-10
: Sequence 10, Application US/09919703
: Patent No. US20020165129A1
: GENERAL INFORMATION:
: APPLICANT: Krysstal, Gerald W.
: APPLICANT: Rabkin, Simon M.
: TITLE OF INVENTION: Peptides and Their Use to Ameliorate
: TITLE OF INVENTION: Cell Death
: FILE REFERENCE: 50216/003004
: CURRENT APPLICATION NUMBER: US/09/919,703
: CURRENT FILING DATE: 2001-07-31
: PRIOR APPLICATION NUMBER: US 09/294,457
: PRIOR FILING DATE: 1999-04-19
: PRIOR APPLICATION NUMBER: US 08/759,599
: PRIOR FILING DATE: 1996-12-05
: PRIOR APPLICATION NUMBER: US 60/008,233
: PRIOR FILING DATE: 1995-12-06
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 10
: LENGTH: 20
: TYPE: PRT
: ORGANISM: Artificial Sequence

```

```

;      FEATURE:
;      OTHER INFORMATION: Synthetic polypeptide
US-09-919-703-10

```

Query Match	8.8%	Score 88;	DB 9;	Length 20;
Best Local Similarity	95.0%	Pred. No. 0.0039;		
Matches 19;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

```
QY 31 RLICDSRVLERYLLLEAKAE 50
    ||| ||||| ||||| |||
Db 1 RLILDSRVLERYLLLEAKAE 20
```

RESULT 14
US-09-729-674-279
Sequence 370 Analysis: US00720674

```

1  APPLICANT:  Jacobs, Kenneth
2  APPLICANT:  McCoy, John M.
3  APPLICANT:  Lavallee, Edward R.
4  APPLICANT:  Collins-Racie, Lisa A.
5  APPLICANT:  Evans, Cheryl
6  APPLICANT:  Merberg, David
7  APPLICANT:  Treacy, Maurice
8  APPLICANT:  Agostino, Michael J.
9  APPLICANT:  Steinger, II, Robert J.
10 APPLICANT:  Spaulding, Vikki
11 APPLICANT:  Wong, Gordon G.
12 APPLICANT:  Clark, Hilary
13 APPLICANT:  Fechtel, Kim
14 APPLICANT:  Genetics Institute, Inc.
15 TITLE OF INVENTION:  SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
16 FILE REFERENCE:  6055-64X
17
18 CURRENT APPLICATION NUMBER:  US/09/729,674
19 CURRENT FILING DATE:  2000-12-04
20 PRIOR APPLICATION NUMBER:  09/559,330
21 PRIOR FILING DATE:  2000-03-30
22 NUMBER OF SEQ. ID NOS:  283
23 SOFTWARE:  Patentl Ver. 2.0
24 SEQ. ID NO. 279

```

Query Match	8.1%	Score 81;	DB 10;	Length 147;
Best Local Similarity	28.8%	Pred. No. 0.32;		
Matches 42;	Conservative 23;	Mismatches 63;	Indels 18;	Gaps 7;

```
Oy 17 LSLPLGLPVLCAAPRRLICDSRVLYRYLLLEAKAEENIT--TCGAEHGS--LNENITVDPDT 71
      :: | | | | | | | | | | | | | | | | | |
Db 1 MASPSGLCVLVRLLPKLLCGGKLTERTLLDLADGTLTKVGVCSSEDAKLLQDGLVVRG 60
```

```

QY      72 KVNFIYAMKREVEVROOAVEVMOGLAL--LSEAVYR--GQALLVNSSQPWEPQLQHYDKAV 126
      ::      :      :      :      :      :      :      :      :      :      :
Db      61 CLDL--RYLAMRQRNNILCNGLSLKSLAEYVLPPLDKSLLLRCSN-WDAETLTEDQVI 116

```

```
Oy 127 SGLR----SLTLLRALGAQKEAISP 148
      |  |  |  |  |  |  |  |  |
Db 117 YARDAQISVALFLHLLGYPFSRNSP 142
```

RESULT 15
 US-09-964-899-25
 Sequence 25, Application US/09964899
 Patent No. US20020174445A1
 GENERAL INFORMATION:
 APPLICANT: Cohen, Dalia et al.
 TITLE OF INVENTION: Identification of Genes Involved in
 TITLE OF INVENTION: Alzheimer's Disease Using Drosophila
 Melanogaster
 FILE REFERENCE: 4-31612 A
 CURRENT APPLICATION NUMBER: US/09/964,899

;; CURRENT FILING DATE: 2001-09-27
;; PRIOR APPLICATION NUMBER: 60/236,893
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/298,309
;; PRIOR FILING DATE: 2001-06-14
;; NUMBER OF SEQ ID NOS: 53
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 25
;; LENGTH: 813
;; TYPE: PRF
;; ORGANISM: Homo Sapien
US-09-964-899-25

Query Match 7.8%; Score 77.5; DB 9; Length 813;
Best Local Similarity 26.5%; Pred. No. 7.8;
Matches 26; Conservative 19; Mismatches 36; Indels 17; Gaps 3;

OY 39 LERYLLEAKEANITTCACPHC-----SLNENITVPPTKVNFYAMKRMVEVROQAVEYMOGL 94
DB 423 LEQQLKEVRRK-----CAEEAQLISSLKRAELISQESQISTYEELAKAREELSRLOQET 476
OY 95 ALISEAVLNGQALLVNSQPMPEPLQLHVDKAVSGLRSL 132
DB 477 AELESVESGKAQL-----EPLQOHLQDSQOEISSM 507

Search completed: January 8, 2003, 05:17:47
Job time : 40.3315 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 8, 2003, 05:15:15 ; Search time 2582.65 Seconds
(without alignments)
2174.835 Million cell updates/sec

Title: US-09-813-775c-2
Perfect score: 998
Sequence: 1 MGVHCCPAMWLLLSLSLP.....NFLRCKLKLYTGACHTGDR 193

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ .p2n.model -DEV=xlh
-O/cqaz2.1/USPTO.spool/US09813775.r/unat.07012003.153122.23818/app.query.fasta_1.718
-DB=GenEmbl -OFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOECL=0 -LOOEEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=Diosun62 -TRANS=human40.cdt1 -LIST=45
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOR=ext -HEA=SITE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09813775.@CN1.1.1741.@unat.07012003.153122.23818 -NCPU=6 -ICPU=3
-MAPN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_jmu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrl:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match		DB ID	Description
		Length	DB		
1	998	100.0	585	6 AX046870	AX046870 Sequence
2	991	99.3	582	6 AX025442	AX025442 Sequence
3	991	99.3	582	6 I05397	I05397 Sequence 2
4	991	99.3	788	6 AR086679	AR086679 Sequence
5	991	99.3	823	6 AR063255	AR063255 Sequence
6	991	99.3	823	6 AR179013	AR179013 Sequence
7	991	99.3	1011	6 A47505	A47505 Sequence 12
8	991	99.3	3100	6 AX451645	AX451645 Sequence
9	983	98.5	1342	6 HSERPR	X02157 Human mRNA
10	955	95.7	724	6 I05399	I05399 Sequence 4
11	899.5	90.1	1344	6 I08348	I08348 Sequence 1
12	899.5	90.1	1462	6 MACRPO	MACRPO
13	899.5	90.1	5107	6 A92667	A92667 Sequence 3
14	899.5	90.1	5107	6 AR158347	AR158347 Sequence
15	894.5	89.6	691	6 MACRXYTHRO	I10609 Macaca mula
16	893.5	89.5	1344	6 I07889	I07889 Sequence 1
17	862	86.4	1292	6 E00630	E00630 DNA encodin
18	849	85.1	514	6 AX150264	AX150264 Sequence
19	843	84.5	501	6 BD007423	BD007423 Cyclicall
20	843	84.5	508	6 I49875	I49875 Sequence 69
21	843	84.5	533	6 I07891	I07891 Sequence 5
22	843	84.5	551	6 I00283	I00283 Sequence 4
23	843	84.5	551	6 I07892	I07892 Sequence 6
24	843	84.5	551	6 I08350	I08350 Sequence 6
25	834	83.6	501	6 AX464420	AX464420 Sequence
26	833	83.3	501	6 AX464421	AX464421 Sequence
27	821	82.3	945	6 A47500	A47500 Sequence 7
28	821	82.3	969	6 A47497	A47497 Sequence 4
29	794.5	79.6	681	4 CATERXYTHRO	CATERXYTHRO
30	793.5	79.5	725	4 FD000685	FD000685 Fells domes
31	780.5	78.2	1263	4 SSC249745	SSC249745 Sus scrofa
32	778.5	78.0	1395	10 RATEPO	RATEPO
33	774.5	77.6	681	4 DOGERYPRE	DOGERYPRE
34	773.5	77.5	585	4 OAEERYPOIA	OAEERYPOIA
35	769.5	77.1	688	4 PIGERYTHRO	PIGERYTHRO
36	767	76.9	679	10 RATERYTHRO	RATERYTHRO
37	766.5	76.8	3609	6 AX249944	AX249944 Sequence
38	761.5	76.3	1133	4 BT044762	BT044762 Bos taurus
39	761.5	76.3	1258	4 BOVEPO	BOVEPO
40	759.5	76.1	689	4 SHPERYTHRO	SHPERYTHRO
41	751.5	75.3	98876	9 AC009488	AC009488 Homo sapi
42	735.5	73.7	3211	6 I07890	I07890 Sequence 3
43	735.5	73.7	3211	6 I08349	I08349 Sequence 3
44	732.5	73.4	3601	6 AR119721	AR119721 Sequence
45	732.5	73.4	3602	6 AR086678	AR086678 Sequence

RESULT 1

ALIGNMENTS

AX046870	585 bp	DNA	linear	PAT 15-DEC-2000
LOCUS				
DEFINITION	Sequence 3 from Patent WO0068376.			
ACCESSION	AX046870			
VERSION	AX046870.1	GI:11876340		
KEYWORDS				
SOURCE				
ORGANISM	chimpanzee, Pan troglodytes			
REFERENCE	Euarctota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.			
AUTHORS	1 (bases 1 to 585)			
TITLE	Desauvage, F. and Hennen, D. J. Chimpanzee erythropoietin (cnepo) polypeptides and nucleic acids encoding the same			
JOURNAL	Patent: WO 0068376-A 3 16-NOV-2000;			
FEATURES	GENENTECH, INC. (US)			
source	location/Qualifiers			
	1..585			
	/organism="Pan troglodytes"			
	/db_xref="taxon:9598"			
BASE COUNT	115 a 185 c 168 g 117 t			
ORIGIN				
Alignment Scores:				
Pred. No.:	2,93e-90	Length:	585	
Score:	998.00	Matches:	193	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	100.00%	Indels:	0	
DB:	6	Gaps:	0	

US-09-813-775C-2 (1-193) x AX046870 (1-585)

QY	1	MetGluValHisGluCysProAlaTyrLeuThrLeuLeuLeuSerLeuLeuSerLeuPro	20
Db	1	ATGGGGGGGCACGAATGCTCGCTGGCTTCTCGTCCGCTGCTGCTCCT	60
QY	21	LeuGlyLeuProValLeuGlyValAProProArgLeuIleCysAspSerArgValLeuGlu	40
Db	61	CTGGGGCTCCACAGTCCCTGGGGCCCCACACAGCCCTCATCTGTGCACACCGAGTCTGGAG	120
QY	41	ArgTyrLeuLeuGluValAlaGlyAlaGlnAsnIleThrThrGlyCysAlaGluHisCys	60
Db	121	AGGTACCTCTTGGAGGGCCAGAGAGCCGAGAAATATCAAGAGGGCTGTGGCAACCTGC	180
QY	61	SerLeuAsnGluAsnIleThrValAProAspThrLysValAsnPhenylAlaTyrLysArg	80
Db	181	AGCTTGAAATGGAATATACACTGCTCCACAGACCAAAATTAATTTCTATGCTTGGAAGAGC	240
QY	81	MetGluValArgGlnGlnAlaValGluValATyrGlnGlyLeuAlaLeuLeuSerGluAla	100
Db	241	ATGGAGGTCAGAGCAGCAGCCCGTGAAGTGTGGCAGGGCCCTGCGCCCTCTCGGAAGCT	300
QY	101	ValLeuAsnArgGlyGlnAlaLeuLeuValAsnSerSerGlnProIrrpGluProLeuGlnLeu	120
Db	301	GTCCGGGGGGCCAGGCCCTGTGGTGTCAACTCTTCCAGCCGCTGGAGCCCTGCACGCTG	360
QY	121	HisValAspLysAlaValAsnSerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGly	140
Db	361	CATGTGGATTAAGCCGTCAGTGGGCTTCGCAGCCTACACACTGCTGCTGGGGCTCTGGGA	420
QY	141	AlaGlnLysGluAlaAlaIleSerProProAspAlaAlaSerAlaAlaProLeuArgThrIle	160
Db	421	GCCCGAAGGAGGCACATCTCCCTCCAGATGGGGCTCACCTGCTCCACTCCGAAACATC	480
QY	161	ThrAlaAspThrPheArgLysLeuPheArgValTyrSerAsnPheLeuAlaGlyLysLeu	180
Db	481	ACTGCTGACACTTTCGCCAAACTTTCGAGTCTATCCCAATTTCCCTCGGGGAAAGCTG	540
QY	181	LysLeuTyrThrGlyLysAlaCysArgThrGlyAspArg	193
Db	541	AAGCTGTACACAGGGGAGGCTCGAGGACAGCAGGGGACAGA	579

LOCUS	AX025442	LOCUS	AX025442	LOCUS	AX025442	LOCUS	AX025442	LOCUS	AX025442		
DEFINITION	Sequence 3 from Patent FR2786104.	DEFINITION	Sequence 3 from Patent FR2786104.	DEFINITION	Sequence 3 from Patent FR2786104.	DEFINITION	Sequence 3 from Patent FR2786104.	DEFINITION	Sequence 3 from Patent FR2786104.		
ACCESSION	AX025442	ACCESSION	AX025442	ACCESSION	AX025442	ACCESSION	AX025442	ACCESSION	AX025442		
VERSION	AX025442.1	VERSION	AX025442.1	VERSION	AX025442.1	VERSION	AX025442.1	VERSION	AX025442.1		
KEYWORDS	GI:10187118	KEYWORDS	GI:10187118	KEYWORDS	GI:10187118	KEYWORDS	GI:10187118	KEYWORDS	GI:10187118		
SOURCE	human.	SOURCE	human.	SOURCE	human.	SOURCE	human.	SOURCE	human.		
ORGANISM	Homo sapiens	ORGANISM	Homo sapiens	ORGANISM	Homo sapiens	ORGANISM	Homo sapiens	ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 582)	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 582)	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 582)	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 582)	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 582)	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 582)
AUTHORS	Haeflner, A. and Hirsch, F.	AUTHORS	Haeflner, A. and Hirsch, F.	AUTHORS	Haeflner, A. and Hirsch, F.	AUTHORS	Haeflner, A. and Hirsch, F.	AUTHORS	Haeflner, A. and Hirsch, F.		
JOURNAL	Patent: FR 2786104-A 3 26-MAY-2000;	JOURNAL	Patent: FR 2786104-A 3 26-MAY-2000;	JOURNAL	Patent: FR 2786104-A 3 26-MAY-2000;	JOURNAL	Patent: FR 2786104-A 3 26-MAY-2000;	JOURNAL	Patent: FR 2786104-A 3 26-MAY-2000;	JOURNAL	Patent: FR 2786104-A 3 26-MAY-2000;
FEATURES	CENTRE NAT RECH SCIENT (FR)	FEATURES	CENTRE NAT RECH SCIENT (FR)	FEATURES	CENTRE NAT RECH SCIENT (FR)	FEATURES	CENTRE NAT RECH SCIENT (FR)	FEATURES	CENTRE NAT RECH SCIENT (FR)		
SOURCE	Location/Qualifiers	SOURCE	Location/Qualifiers	SOURCE	Location/Qualifiers	SOURCE	Location/Qualifiers	SOURCE	Location/Qualifiers		
	1..582		1..582		1..582		1..582		1..582		
	/organism="Homo sapiens"		/organism="Homo sapiens"		/organism="Homo sapiens"		/organism="Homo sapiens"		/organism="Homo sapiens"		
	/db_xref="taxon:9606"		/db_xref="taxon:9606"		/db_xref="taxon:9606"		/db_xref="taxon:9606"		/db_xref="taxon:9606"		
	1..582		1..582		1..582		1..582		1..582		
	/note="unnamed protein product"		/note="unnamed protein product"		/note="unnamed protein product"		/note="unnamed protein product"		/note="unnamed protein product"		
	/codon_start=1		/codon_start=1		/codon_start=1		/codon_start=1		/codon_start=1		
	/protein_id="CAC09044.1"		/protein_id="CAC09044.1"		/protein_id="CAC09044.1"		/protein_id="CAC09044.1"		/protein_id="CAC09044.1"		
	/db_xref="GI:10187119"		/db_xref="GI:10187119"		/db_xref="GI:10187119"		/db_xref="GI:10187119"		/db_xref="GI:10187119"		
	/translation="MGVHECPAMTLLSLSLPLGIPVIGAPPRLICDSNVLERYLL EAKRAENITTCGAHCISLNENITVPDKVNFYARKRAEVEGQAAVEWGGALLSEAVY RGQALLNNSQPEPLQIHDVKAVSGLRSLTLLRALGAQKAEATSPDDAASAPLRTIT TADPFRKLFNYSNPLRGKLTLYTGCACRCGDR"		/translation="MGVHECPAMTLLSLSLPLGIPVIGAPPRLICDSNVLERYLL EAKRAENITTCGAHCISLNENITVPDKVNFYARKRAEVEGQAAVEWGGALLSEAVY RGQALLNNSQPEPLQIHDVKAVSGLRSLTLLRALGAQKAEATSPDDAASAPLRTIT TADPFRKLFNYSNPLRGKLTLYTGCACRCGDR"		/translation="MGVHECPAMTLLSLSLPLGIPVIGAPPRLICDSNVLERYLL EAKRAENITTCGAHCISLNENITVPDKVNFYARKRAEVEGQAAVEWGGALLSEAVY RGQALLNNSQPEPLQIHDVKAVSGLRSLTLLRALGAQKAEATSPDDAASAPLRTIT TADPFRKLFNYSNPLRGKLTLYTGCACRCGDR"		/translation="MGVHECPAMTLLSLSLPLGIPVIGAPPRLICDSNVLERYLL EAKRAENITTCGAHCISLNENITVPDKVNFYARKRAEVEGQAAVEWGGALLSEAVY RGQALLNNSQPEPLQIHDVKAVSGLRSLTLLRALGAQKAEATSPDDAASAPLRTIT TADPFRKLFNYSNPLRGKLTLYTGCACRCGDR"		/translation="MGVHECPAMTLLSLSLPLGIPVIGAPPRLICDSNVLERYLL EAKRAENITTCGAHCISLNENITVPDKVNFYARKRAEVEGQAAVEWGGALLSEAVY RGQALLNNSQPEPLQIHDVKAVSGLRSLTLLRALGAQKAEATSPDDAASAPLRTIT TADPFRKLFNYSNPLRGKLTLYTGCACRCGDR"		
BASE COUNT	113 a	BASE COUNT	113 a	BASE COUNT	113 a	BASE COUNT	113 a	BASE COUNT	113 a		
ORIGIN	181 c	ORIGIN	181 c	ORIGIN	181 c	ORIGIN	181 c	ORIGIN	181 c		
	170 g		170 g		170 g		170 g		170 g		
	118 t		118 t		118 t		118 t		118 t		

Alignment Scores:	1,45e-89	Length:	582
Score:	991.00	Matches:	192
Percent Similarity:	99.48%	Conservative:	0
Best Local Similarity:	99.48%	Mismatch:	1
Query Match:	99.30%	Indels:	0
DB:	6	Gaps:	0

US-09-813-775C-2 (1-193) x AX025442 (1-582)

QY	1	MetGlyValHisGluCysProAlaTrpLeuTrpLeuLeuSerLeuLeuSerLeuPro	20
DB	1	ATGGGGGGGCACGAATGCTGCTGGCTGTGGCTTCTCTGCTCCCTGCTGGCTCCCT	60
QY	21	LeuGlyLeuProValLeuGlyAlaProProArgLeuIleCysAspSerArgValLeuGlu	40
DB	61	CTGGGCTCCCGACGCTCTGGGGGGCCCGCCACGCTCATCTGTGACAGCCGAGATCTGGAG	120
QY	41	ArgTyrLeuLeuGluValAlaLysGluValGluLysAsnIleThrThyGlyCysAlaGluHisCys	60
DB	121	AGGTACCTCTTGGAGAGCCACAGAGCCGAGATATCACAGACGGCTGTGTCAACACTGC	180
QY	61	SerLeuAsnGluAsnIleThrValProAspThrIysValAsnPheTyrAlaTrpLysArg	80
DB	181	AGCTTGAAATGGAATATCTACTGTCCCAACACCAAAAGTTAATTTCTATGGCTGGAGAGAG	240
QY	81	MetGluValArgGlnGlnAlaValGluValArgGlnGlyLeuAlaLeuLeuSerLeuAla	100
DB	241	ATGGAGGTGGGGACACGACGCGGTGAAGTCTGGGACGGGCTGGCCCTCTGTGGAGCT	300
QY	101	ValLeuArgGlyGlnAlaLeuLeuValAsnSerSerGlnProTrpGluProLeuGlnLeu	120
DB	301	GTCCGTGGGGGGACAGGCCCTGTGTGTCAACTTTCACAGCGGTGGAGCCCTTCGACGTG	360
QY	121	HisValAspLysAlaValSerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGly	140
DB	361	CATGTGGATTAAGCCGTCAGTGGCTTGTGCAGACCTCACACTCTCTCTGGGCTTGGGA	420
QY	141	AlaGlnLysGluAlaIleSerProProAspAlaAlaSerAlaAlaProLeuArgThrIle	160
DB	421	GCCCAAGAGAACCATCTCCCTCCCAATGAGGCGCTAGCTGTCCACTCCGAACATTC	480

1

QY 161 ThrAlaasphrPhearglyLeupheargValIyrSeranpheLeuArgIlyLysleu 180
|||||
Db 481 ACTGCTACACTTCCCAAACTCTCGAGTACTCAATTCTCCGGGAAAGCTG 540

QY 181 LysleuTyrrhrgIygluaIaCysargThrgIyAsparg 193
|||||
Db 541 AAGCTGTACACAGGAGGAGGCTGCAGACAGGAGGACAGA 579

RESULT 3
LOCUS 105397 582 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 2 from Patent EP 0267678.
ACCESSION 105397
VERSION 105397.1 GI:590970
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS Beck,A.K., Withy,R.M., Zabrecky,J.R. and Mastello,N.C.
TITLE Recombinant human erythropoietin
JOURNAL Patent: EP 0267678-A1 2.18-MAY-1988:
FEATURES
source location/Qualifiers
1..582
/organism="unknown"

BASE COUNT 113 a 181 c 170 g 118 t
ORIGIN

Alignment Scores:
Pred. No.: 1,45e-89 Length: 582
Score: 991.00 Matches: 192
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 99.30% Indels: 0
gaps: 0

US-09-813-775c-2 (1-193) x 105397 (1-582)

QY 1 MetGlyValHisGluCysProAlaTrpLeuTrpLeuLeuSerLeuLeuSerLeuPro 20
|||||
Db 1 ATGGGGGTGCAGCAATGCTCGCTGGCTTCTCCCTGCTGCTGCTGCTGCTGCT 60

QY 21 LeuGlyLeuProValLeuGlyAlaProProArgLeuIleCysAspSerArgValLeuGlu 40
|||||
Db 61 CTGGGCTCCAGTCCTGGGGGCCCAACAGCCTCATCTGTGACAGCGAGTCTGGAG 120

QY 41 ArgTyrrLeuLeuGluAlaIalysGluAlaGluAsnIleThrrhrgIyCysAlaGluHisCys 60
|||||
Db 121 AGTACCTCTTGGAGGCCAAGAGGCCAGCAATATCACAGAGGCGCTGTGAACTGTC 180

QY 61 SerLeuAsnGluAsnIleThrrValProAspThrrLysValAsnPhetyrAlaTrpLysArg 80
|||||
Db 181 AGCTTGAATGAGATATACAGTCTCCAGACCAAAAGTTATTCTATGCTGGAAAGAG 240

QY 81 MetGluValArgGlnGlnAlaValGluValTrpGlnGlyLeuAlaLeuLeuSerGluAla 100
|||||
Db 241 ATGAGGTCTGGGCGAGCAGCGGTAGAAATCTGGAGGCGCTGGCTGTGCGAAAGCT 300

QY 101 ValLeuArgGlyGlnAlaLeuLeuValAsnSerSerGlnProTrpGluProLeuGlnLeu 120
|||||
Db 301 GTCTGCGGGGCCAGGCGCTGTGTGTCAACTTCTCCAGCGCTGGAGCGCTGTCAGCTG 360

QY 121 HisValAspLysAlaValSerGlyLeuArgSerLeuThrrhrgLeuLeuArgAlaLeuGly 140
|||||
Db 361 CATGTGATTAAGCGCTCAGTGGCTTGCAGCTCACCACTGTGCTCGGGCTGTGGGA 420

QY 141 AlaGlnLysGluAlaIalSerProProAspAlaIalSerAlaIalProLeuArgThrIle 160
|||||
Db 421 GCCCAGAAGGAAGCAATCTCCCTCCAGATCGGCGCTCAGCTGCCACTCCGAACAATC 480

QY 161 ThrAlaAspThrrPheargLysLeuPheargValIyrSeranpheLeuArgIlyLysleu 180
|||||
Db 481 ACTGCTACACTTCCCAAACTCTCGAGTACTCAATTCTCCGGGAAAGCTG 540

QY 181 LysleuTyrrhrgIygluaIaCysargThrgIyAsparg 193
|||||
Db 541 AAGCTGTACACAGGAGGAGGCTGCAGACAGGAGGACAGA 579

RESULT 4
LOCUS AR086679 788 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 35 from patent US 5985607.
ACCESSION AR086679
VERSION AR086679.1 GI:10013445
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS Delcuve,G. and Awang,G.
TITLE Recombinant DNA molecules and expression vectors for tissue plasminogen activator
JOURNAL Patent: US 5985607-A 35.16-NOV-1999:
FEATURES
source location/Qualifiers
1..788
/organism="unknown"

BASE COUNT 145 a 264 c 228 g 151 t
ORIGIN

Alignment Scores:
Pred. No.: 2,06e-89 Length: 788
Score: 991.00 Matches: 192
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 99.30% Indels: 0
gaps: 0

US-09-813-775c-2 (1-193) x AR086679 (1-788)

QY 1 MetGlyValHisGluCysProAlaTrpLeuTrpLeuLeuSerLeuLeuSerLeuPro 20
|||||
Db 80 ATGGGGGTGCAGCAATGCTCGCTGGCTTCTCCCTGCTGCTGCTGCTGCTGCT 139

QY 21 LeuGlyLeuProValLeuGlyAlaProProArgLeuIleCysAspSerArgValLeuGlu 40
|||||
Db 140 CTGGGCTCCAGTCCTGGGGGCCCAACAGCCTCATCTGTGACAGCGAGTCTGGAG 199

QY 41 ArgTyrrLeuLeuGluAlaIalysGluAlaGluAsnIleThrrhrgIyCysAlaGluHisCys 60
|||||
Db 200 AGTACCTCTTGGAGGCCAAGAGGCCAGCAATATCACAGAGGCGCTGTGAACTTGC 259

QY 61 SerLeuAsnGluAsnIleThrrValProAspThrrLysValAsnPhetyrAlaTrpLysArg 80
|||||
Db 260 AGCTTGAATGAGATATACAGTCTCCAGACCAAAAGTTATTCTATGCTGGAAAGAG 319

QY 81 MetGluValArgGlnGlnAlaValGluValTrpGlnGlyLeuAlaLeuLeuSerGluAla 100
|||||
Db 320 ATGAGGTCTGGGCGAGCAGCGGTAGAAATCTGGAGGCGCTGGCTGTGCGAAAGCT 379

QY 101 ValLeuArgGlyGlnAlaLeuLeuValAsnSerSerGlnProTrpGluProLeuGlnLeu 120
|||||
Db 380 GTCTGCGGGGCCAGGCGCTGTGTGTCAACTTCTCCAGCGCTGGAGCGCTGTCAGCTG 439

QY 121 HisValAspLysAlaValSerGlyLeuArgSerLeuThrrhrgLeuLeuArgAlaLeuGly 140
|||||
Db 440 CATGTGATTAAGCGCTCAGTGGCTTGCAGCTCACCACTGTGCTCGGGCTGTGGGA 499

QY 141 AlaGlnLysGluAlaIalSerProProAspAlaIalSerAlaIalProLeuArgThrIle 160
|||||
Db 500 GCCCAGAAGGAAGCAATCTCCCTCCAGATCGGCGCTCAGCTGCCACTCCGAACAATC 559

QY 161 ThrAlaAspThrrPheargLysLeuPheargValIyrSeranpheLeuArgIlyLysleu 180
|||||
Db 560 ACTGCTACACTTCCCAAACTCTCGAGTACTCAATTCTCCGGGAAAGCTG 619

QY 181 LysleuTyrrhrgIygluaIaCysargThrgIyAsparg 193

Db 620 AAGCTGTACACAGGAGGCGCTGCAGACAGGAGACAGA 658
|||||
RESULT 5
LOCUS AR063255 823 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5846528.
ACCESSION AR063255
VERSION AR063255.1 GI:5992563
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE
AUTHORS 1 (bases 1 to 823)
TITLE Podsakoff, G.M. and Kurtzman, G.J.
JOURNAL Treating anemia using recombinant adeno-associated virus virions
FEATURES
source Patent: US 5846528-A 1 08-DEC-1998;
location/Qualifiers
BASE COUNT 159 a 275 c 233 g 156 t
ORIGIN
Alignment Scores:
Pred. No.: 2,16e-89 Length: 823
Score: 991.00 Matches: 192
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 99.30% Indels: 0
Gaps: 0
DB: 6
US-09-813-775c-2 (1-193) x AR063255 (1-823)
Qy 1 MetGlyValHisGluCysProAlaTrpLeuTrpLeuLeuSerLeuLeuSerLeuPro 20
Db 52 ATGGGGGTCACAGAAATGCTCTGCGTGGCTTCTCTGCTGCTGCTGCTGCTGCT 111
Qy 21 LeuGlyLeuProValLeuGlyAlaProProAlaGlyLeuIleCysAspSerArgValLeuGlu 40
Db 112 CTGGGCTCTCCAGTCTCTGGGCGCCACACACGCTCATCTGTGACAGCGGAGCTCTGGAG 171
Qy 41 ArgTyrlLeuGluGluAlaValSerGluAlaGluAsnIleThrThrGlyCysAlaGluHisCys 60
Db 172 AGGTACTCTTGGAGGCCAAGAGGCGCGAATATCAGACAGGCGCTGTGTAACACTGC 231
Qy 61 SerLeuAsnGluAsnIleThrValProAspThrLysValAsnPhetYrAlaTrpLysArg 80
Db 232 ACCTGAATGAGAAATATCATCTGCCAGACACCAAAATTAATTCTATGCTCGGAAGAG 291
Qy 81 MetGluValArgGlnGlnAlaValGluValTrpGlnGlyLeuAlaLeuSerGluAla 100
Db 292 ATGGAGTCTGGGACAGGCGCTAGAACTGTGGAGGCGCTGCGCTGTGCGAAGCT 351
Qy 101 ValLeuArgGlyGluAlaLeuLeuValAsnSerSerGlnProTrpGluProLeuGlnLeu 120
Db 352 GTCTGCGGGGCGGAGCCCTGTTGGTCAACTTCTCCACGCGTGGAGGCCCTGCAGCTG 411
Qy 121 HisValAspLysAlaValSerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGly 140
Db 412 CATGTGGAATAAAGCCGTCAGTGGCTTCGCACCTCACCACACTGCTTGGGCTCGGGA 471
Qy 141 AlaGlnLysGluAlaIleSerProProAspAlaAlaSerAlaAlaProLeuArgThrIle 160
Db 472 GCCCAAGAGGAAGCATCTCCCTCCAGATGCGGCGTCAAGTGTCCATCCGAAACAATC 531
Qy 161 ThrAlaAspThrPheArgLysLeuPheArgValIlyrSerAsnPhetLeuArgLysLysLeu 180
Db 532 ACTGTGACACTTTCGCGAAACTCTTCGAGTCTACTCCAAATTCCTCCGCGGGAAGCTG 591
Qy 181 LysLeuTyrlThrGlyGluAlaCysArgTrpTrgIysAspArg 193
Db 592 AAGCTGTACACAGGAGGCGCTGCAGACAGGAGACAGA 630
|||||
RESULT 7

Db 620 AAGCTGTACACAGGAGGCGCTGCAGACAGGAGACAGA 658
|||||
RESULT 6
LOCUS AR179013 823 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 1 from patent US 6325998.
ACCESSION AR179013
VERSION AR179013.1 GI:20220568
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE
AUTHORS 1 (bases 1 to 823)
TITLE Podsakoff, G.M. and Kurtzman, G.J.
JOURNAL Methods of treating disease using recombinant adeno-associated virus virions administered to muscle
FEATURES
source Patent: US 6325998-A 1 04-DEC-2001;
location/Qualifiers
BASE COUNT 159 a 275 c 233 g 156 t
ORIGIN
Alignment Scores:
Pred. No.: 2,16e-89 Length: 823
Score: 991.00 Matches: 192
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 99.30% Indels: 0
Gaps: 0
DB: 6
US-09-813-775c-2 (1-193) x AR179013 (1-823)
Qy 1 MetGlyValHisGluCysProAlaTrpLeuTrpLeuLeuSerLeuLeuSerLeuPro 20
Db 52 ATGGGGGTCACAGAAATGCTCTGCGTGGCTTCTCTGCTGCTGCTGCTGCTGCT 111
Qy 21 LeuGlyLeuProValLeuGlyAlaProProAlaGlyLeuIleCysAspSerArgValLeuGlu 40
Db 112 CTGGGCTCTCCAGTCTCTGGGCGCCACACACGCTCATCTGTGACAGCGGAGCTCTGGAG 171
Qy 41 ArgTyrlLeuGluGluAlaValSerGluAlaGluAsnIleThrThrGlyCysAlaGluHisCys 60
Db 172 AGGTACTCTTGGAGGCCAAGAGGCGCGAATATCAGACAGGCGCTGTGTAACACTGC 231
Qy 61 SerLeuAsnGluAsnIleThrValProAspThrLysValAsnPhetYrAlaTrpLysArg 80
Db 232 ACCTGAATGAGAAATATCATCTGCCAGACACCAAAATTAATTCTATGCTCGGAAGAG 291
Qy 81 MetGluValArgGlnGlnAlaValGluValTrpGlnGlyLeuAlaLeuSerGluAla 100
Db 292 ATGGAGTCTGGGACAGGCGCTAGAACTGTGGAGGCGCTGCGCTGTGCGAAGCT 351
Qy 101 ValLeuArgGlyGluAlaLeuLeuValAsnSerSerGlnProTrpGluProLeuGlnLeu 120
Db 352 GTCTGCGGGGCGGAGCCCTGTTGGTCAACTTCTCCACGCGTGGAGGCCCTGCAGCTG 411
Qy 121 HisValAspLysAlaValSerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGly 140
Db 412 CATGTGGAATAAAGCCGTCAGTGGCTTCGCACCTCACCACACTGCTTGGGCTCGGGA 471
Qy 141 AlaGlnLysGluAlaIleSerProProAspAlaAlaSerAlaAlaProLeuArgThrIle 160
Db 472 GCCCAAGAGGAAGCATCTCCCTCCAGATGCGGCGTCAAGTGTCCATCCGAAACAATC 531
Qy 161 ThrAlaAspThrPheArgLysLeuPheArgValIlyrSerAsnPhetLeuArgLysLysLeu 180
Db 532 ACTGTGACACTTTCGCGAAACTCTTCGAGTCTACTCCAAATTCCTCCGCGGGAAGCTG 591
Qy 181 LysLeuTyrlThrGlyGluAlaCysArgTrpTrgIysAspArg 193
Db 592 AAGCTGTACACAGGAGGCGCTGCAGACAGGAGACAGA 630
|||||
RESULT 7

A47505
LOCUS A47505 1011 bp DNA linear PAT 07-MAR-1997
DEFINITION Sequence 12 from Patent WO9533057.
ACCESSION A47505
VERSION A47505.1 GI:2301471
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
BASE COUNT
ORIGIN

1 (bases 1 to 1011)
Mele, A., De, S. R., Carloni, C. and Coscarella, A.
HEMATOPOIETIC STIMULATION
Patent: WO 9533057-A 12 07-DEC-1995;
MENARINI RICERCHE SUD SPA (IT)
Other publication AU 2616895 951221.
location/Qualifiers
1..1011
/organism="unidentified"
/db_xref="taxon:32644"

213 a 318 c 291 g 189 t

Alignment Scores:
Pred. No.: 2,75e-89 Length: 1011
Score: 991.00 Matches: 192
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 99.30% Indels: 0
DB: 6 Gaps: 0

US-09-813-775C-2 (1-193) x A47505 (1-1011)

QY 1 MetGlyValHisGlyCysProAlaTrpLeuTrpLeuLeuSerLeuLeuSerLeuPro 20
DB 1 ATGGGGGTGCACGAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
QY 21 LeuGlyLeuProValLeuGlyAlaProProArgLeuIleCysAspSerArgValLeuGlu 40
DB 61 CTGGGCTCCAGTCCTGGGCGGCGCCACACAGCCTCATCTGTGACAGCGAGCTCTGAG 120
QY 41 ArgTyrLeuLeuGluAlaValGlyAlaGluAsnIleThrThrGlyCysAlaGluHisCys 60
DB 121 AGGTACTCTTGGAGGCGCAAGAGCGGAGAAATATACAGCGGGCTGTGCTGAACACTGC 180
QY 61 SerLeuAsnGluAsnIleThrValProAspThrLysValAsnPheTyrAlaTrpLysArg 80
DB 181 AGCTTGATGAGATATACATCTGCTCCGACACCAAGTTATTTCTATGCTGGAAAGAG 240
QY 81 MetGluValArgGlnGlnAlaValGluValTrpGlnGlyLeuAlaLeuLeuSerGluAla 100
DB 241 ATGAGAGTCCGCGCAGCAGCGCGTAGAGTGTGGCAGGCGCTGGCCCTGCTCGGAAGCT 300
QY 101 ValLeuArgGlyGlnAlaLeuLeuValAsnSerSerGlnProTrpGluProLeuGlnLeu 120
DB 301 GTCTCTGGGGGCGCAGGCGCTGTGGTCACTCTTCCAGCGGTGGAGCCCTTCACACTG 360
QY 121 HisValAspLysAlaValSerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGly 140
DB 361 CATGTGATTAAGCGCGTCAATGTGGCTTCGCAAGCTTCACACTGCTTGGGGCTCTGGGA 420
QY 141 AlaGlnGlyGluAlaIleSerProProAspAlaIleSerAlaAlaProLeuArgThrIle 160
DB 421 GCCCAGAAAGAGGCAATCTCCCTCCAGATGGCGCTCAGCTGCTCCACTCCGAAACATC 480
QY 161 ThrAlaAspThrPheArgLysLeuPheArgValTyrSerAsnPheLeuArgGlyLysLeu 180
DB 481 ACTGCTGACACTTTCGCAAACTCTTCCGAGTCTACTCCAAATTTCTCTCGGGGAAAGCTG 540
QY 181 LysLeuTyrThrGlyGluAlaCysArgThrGlyAspArg 193
DB 541 AAGCTGTACACAGGGAGGCTGACAGACAGGGAGACAGA 579

RESULT 8
AX451645
LOCUS AX451645 3100 bp DNA linear PAT 03-JUL-2002
DEFINITION Sequence 23 from Patent WO0224899.
ACCESSION AX451645
VERSION AX451645.1 GI:21698583
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
BASE COUNT
ORIGIN

1
Abruzzese, R. V., Mehta, V. and Nordstrom, J. L.
Improved system for regulation of transgene expression
Patent: WO 0224899-A 23 28-MAR-2002;
Valentis Inc. (US)
location/Qualifiers
1..3100
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Complete nucleic acid sequence of pEP1666 plasmid."

656 a 913 c 856 g 675 t

Alignment Scores:
Pred. No.: 1.01e-88 Length: 3100
Score: 991.00 Matches: 192
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 99.30% Indels: 0
DB: 6 Gaps: 0

US-09-813-775C-2 (1-193) x AX451645 (1-3100)

QY 1 MetGlyValHisGlyCysProAlaTrpLeuTrpLeuLeuSerLeuLeuSerLeuPro 20
DB 417 ATGGGGGTGCACGAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 476
QY 21 LeuGlyLeuProValLeuGlyAlaProProArgLeuIleCysAspSerArgValLeuGlu 40
DB 477 CTGGGCTCCAGTCCTGGGCGGCGCCACACAGCCTCATCTGTGACAGCGGCTCTGAG 536
QY 41 ArgTyrLeuLeuGluAlaValGlyAlaGluAsnIleThrThrGlyCysAlaGluHisCys 60
DB 537 AGGTACTCTTGGAGGCGCAAGAGCGGAGAAATATACAGCGGGCTGTGCTGAACACTGC 596
QY 61 SerLeuAsnGluAsnIleThrValProAspThrLysValAsnPheTyrAlaTrpLysArg 80
DB 597 AGCTTGATGAGATATACATCTGCTCCGACACCAAGTTATTTCTATGCTGGAAAGAG 656
QY 81 MetGluValArgGlnGlnAlaValGluValTrpGlnGlyLeuAlaLeuLeuSerGluAla 100
DB 657 ATGAGAGTCCGCGCAGCAGCGCGTAGAGTGTGGCAGGCGCTGGCCCTGCTCGGAAGCT 716
QY 101 ValLeuArgGlyGlnAlaLeuLeuValAsnSerSerGlnProTrpGluProLeuGlnLeu 120
DB 717 GTCTCTGGGGGCGCAGGCGCTGTGGTCACTCTTCCAGCGGTGGAGCCCTTCACACTG 776
QY 121 HisValAspLysAlaValSerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGly 140
DB 777 CATGTGATTAAGCGCGTCAATGTGGCTTCGCAAGCTTCACACTGCTTGGGGCTCTGGGA 836
QY 141 AlaGlnGlyGluAlaIleSerProProAspAlaIleSerAlaAlaProLeuArgThrIle 160
DB 837 GCCCAGAAAGAGGCAATCTCCCTCCAGATGGCGCTCAGCTGCTCCACTCCGAAACATC 896
QY 161 ThrAlaAspThrPheArgLysLeuPheArgValTyrSerAsnPheLeuArgGlyLysLeu 180
DB 897 ACTGCTGACACTTTCGCAAACTCTTCCGAGTCTACTCCAAATTTCTCTCGGGGAAAGCTG 956
QY 181 LysLeuTyrThrGlyGluAlaCysArgThrGlyAspArg 193
DB 957 AAGCTGTACACAGGGAGGCTGACAGACAGGGAGACAGA 995

RESULT 9
 HSERPR 1342 bp mRNA linear PRI 30-MAR-1995
 LOCUS Human mRNA for fetal erythropoietin.
 DEFINITION X02157
 ACCESSION X02157
 VERSION X02157.1 GI:31229
 KEYWORDS erythropoietin; glycoprotein hormone; hormone; signal peptide.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 1342)
 AUTHORS Jacobs, K., Shoemaker, C., Rudersdorf, R., Neill, S.D., Kaufman, R.J.,
 Mulson, A., Seehra, J., Jones, S.S., Hewick, R., Fritsch, E.F.,
 Kawakita, M., Shimizu, T. and Miyake, T.
 TITLE Isolation and characterization of genomic and cDNA clones of human
 erythropoietin
 JOURNAL Nature 313 (6005), 806-810 (1985)
 MEDLINE 85137899
 PUBMED 3838366
 COMMENT Data kindly reviewed (24-FEB-1986) by K. Jacobs.
 FEATURES
 location/Qualifiers
 source 1..1342
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 CDS 182..763
 /note="erythropoietin precursor"
 /codon_start=1
 /protein_id="CAA26094.1"
 /db_xref="GI:31230"
 /db_xref="SWISS-PROT:P01588"
 /translation="MGVHECPAMWLLSLSLPLGLFVLCAPPRLTDSRYLERYLL
 EAKEENITSCAEHCSLNNITVPDITVNVYAMRMVGGQAVVWGLALSEAVL
 RGOALVNSQPEPLQLHVDKAVSGLSLFTLLALRAKREKAISSPPDAASAPLRTI
 TADPRKLFRRYSNPLRGLKLYGCACTGDR"
 misc_feature 182..762
 /product="erythropoietin (aa 1-166)"
 misc_feature 332..334
 /note="glycosylation site"
 misc_feature 374..376
 /note="glycosylation site"
 misc_feature 509..511
 /note="glycosylation site"
 BASE COUNT 276 a 434 c 385 g 247 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2,388-88 Length: 1342
 Score: 983.00 Matches: 191
 Percent Similarity: 98.96% Conservative: 0
 Best Local Similarity: 98.96% Mismatches: 2
 Query Match: 98.50% Indels: 0
 DB: 9 Gaps: 0
 US-09-813-775c-2 (1-193) x HSERPR (1-1342)
 QY 1 MetGlyValHisGluCysProAlaTrpLeuTrpLeuLeuSerLeuLeuSerLeuPro 20
 Db 182 ATGGGGGTCACGAAATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241
 QY 21 LeuGlyLeuProValLeuGlyAlaProProAlaGlyLeuIleCysAspSerArgValLeuGlu 40
 Db 242 CTGGGCTCTCCAGCTCTGGGCGCCACACGCCCTCATCTGTGACAGCGCGCTCTGGAG 301
 QY 41 ArgTyrLeuLeuGluAlaValGluAlaGluAsnIleThrThyGlyCysAlaGluHisCys 60
 Db 302 AGGTACCTCTTGGAGGCCAAGAGAGGCCAGAAATATCAGACGAGGCTGTCTGAACACTGC 361
 QY 61 SerLeuAsnGluAsnIleThrValProAspThrIleCysValAsnPheTyrAlaTrpLysArg 80
 Db 362 AGCTTGAAATGAGAAATATACCTGCTCCACAGACCAAGTAATTCTATGCTCTGGAAGAGG 421

QY 81 MetGlyValArgGlnGlnAlaValGluValTrpGlnGlyLeuAlaLeuLeuSerGlnAla 100
 Db 422 ATGAGGTCTGGCGACAGCCCGTAGAAGTCTGGCAGGCGCTGCTGCTGCGAAGCT 481
 QY 101 ValGlnArgGlyGlnAlaLeuLeuValAsnSerSerGlnProTrpGlnProLeuGlnLeu 120
 Db 482 GTCTCTGGGGGCCAGGCGCTGTGTGTCAACTCTCCACCGCTGGAGGCCCTGCACTG 541
 QY 121 HisValAspLysAlaValSerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGly 140
 Db 542 CATGTGATTAAGCCGTCATGTGGCTTCGACGCTCAGCAGCTCTGCTGGCTCTGCGCA 601
 QY 141 AlaGlnLysGluAlaIleSerProProAspAlaIleSerAlaAlaProLeuArgThrIle 160
 Db 602 GCCCAAGAGGAGCCATCTCCCTCCAGATGGCGCTCAGCTCCAGCTCCAGAACAAATC 661
 QY 161 ThrAlaAspThrPheArgLysLeuPheArgValTyrSerAsnPheLeuArgGlyLysLeu 180
 Db 662 ACTGCTGACACTTTCGCAACTCTTCGAGTCTACTCCAAATTCCTCCGGGAAAGCTG 721
 QY 181 LysLeuTyrThrGlyGluAlaCysArgThrGlyAspArg 193
 Db 722 AAGCTGTACACAGGGAGGCGCTGCAGAGACAGGAGGAGACAGA 760

RESULT 10
 LOCUS 105399 724 bp DNA linear PAT 02-DEC-1994
 DEFINITION Sequence 4 from Patent EP 0267678.
 ACCESSION 105399
 VERSION I05399.1 GI:590971
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 724)
 AUTHORS Beck, A.K., Muthy, R.M., Zabrecky, J.R. and Mastello, N.C.
 TITLE Recombinant human erythropoietin
 JOURNAL Patent: EP 0267678-A1 4 18-MAY-1988;
 FEATURES location/Qualifiers
 source 1..724
 /organism="unknown"
 BASE COUNT 141 a 235 c 202 g 146 t
 ORIGIN

Alignment Scores:
 Pred. No.: 7,066-86 Length: 724
 Score: 955.00 Matches: 187
 Percent Similarity: 98.94% Conservative: 0
 Best Local Similarity: 98.94% Mismatches: 2
 Query Match: 95.69% Indels: 0
 DB: 6 Gaps: 0
 US-09-813-775c-2 (1-193) x I05399 (1-724)
 QY 5 GluCysProAlaTrpLeuTrpLeuLeuSerLeuLeuSerLeuProLeuGlyLeuPro 24
 Db 62 GAATGCTCTGCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 121
 QY 25 ValLeuGlyAlaProProArgLeuIleCysAspSerArgValLeuGluArgTyrLeuLeu 44
 Db 122 GTCTCTGGGGGCCACACAGCTCTCATCTGTGACAGCCAGCTCTGGAGAGGTACTCTTG 181
 QY 45 GluAlaLysGluAlaGluAsnIleThrThrGlyCysAlaGlnHisCysSerLeuAsnGlu 64
 Db 182 GAGGCCAAGAGAGCGAGAAATATCAAGAGGCGCTGTGCGAACACAGCTTGATGAG 241
 QY 65 AsnIleThrValProAspThrIleCysValAsnPheTyrAlaTrpLysArgMetGluValArg 84
 Db 242 AATATCACTGTCTCCAGACACCAAGCTTAATTCTATGCTCTGGAAGAGATGAGGTCGGG 301
 QY 85 GlnGlnAlaValGluValTrpGlnGlyLeuAlaLeuLeuSerGlnAlaValLeuArgGly 104
 Db 302 CAGCAGGCGGTAGAAAGTCTGGCAGGCGCTGCTGTCTGTGGAAGCTGTCTCTGGGGGC 361

[illegible]

QY	121	HISValAspIysAlaValSerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGly	140
Db	500	CACATGGATGAAGGACATGTCGCTTCGGACACACACACTCTGCTCGGGCGCTGGGA	559
QY	141	AAAGlnIysGluAlaIleSerProProAspAlaAlaSerAlaAlaProLeuArgThrIle	160
Db	560	GGCCGAG---GAAGGCATCTCCCTCCAGATGGCGGCTCGGCTGCTCCACTCGAAGCATC	616
QY	161	ThrAlaAspThrPheArgIysLeuPheArgValTyrSerAsnPheLeuArgIlyIysLeu	180
Db	617	ACTGCTGACACTTCTGGCAAACTCTTCGAGCTCACTCAATTCTCCGGGAAAGCTG	676
QY	181	LysLeuTyrThrGlyGluAlaCysArgThrGlyAspArg	193
Db	677	AAGCTGTACACGGGAGGCGCTGCAGAGAGAGGAGGAGCA	715
RESULT 12			
LOCUS	MACEPO	1462 bp	linear PRI 04-AUG-1993
DEFINITION	Monkey (Cynomolgus) erythropoietin mRNA, complete cds.		
ACCESSION	M18189	M15818	M18188
VERSION	M18189.1	GI:342093	
KEYWORDS	erythropoietin.		
SOURCE	Monkey cDNA to mRNA.		
ORGANISM	Macaca fascicularis		
REFERENCE	Umarjota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.		
AUTHORS	Lin,F.K., Lin,C.H., Lai,P.H., Browne,J.K., Eagle,J.C., Smalling,R., Fox,G.M., Chen,K.K., Castro,M. and Suggs,S.		
TITLE	Monkey erythropoietin gene: cloning, expression and comparison with the human erythropoietin gene		
JOURNAL	Gene 44 (2-3), 201-209 (1986)		
MEDLINE	87055236		
PUBMED	2877922		
COMMENT	Drift entry and computer-readable sequence for [1] kindly provided by F. K.Lin, 04-AUG-1987.		
FEATURES	Location/Qualifiers		
source	1..1462		
CDS	/organism="Macaca fascicularis"		
	/db_xref="taxon:9541"		
	217..795		
	/note="erythropoietin precursor"		
	/codon_start=1		
	/protein_id="AAA36841.1"		
	/db_xref="GI:342094"		
	/translation="MGVHECPAMLLLSVSLPLGVPVGPAPRLICDSRVLERYLL		
	EAKAEAVNTMGCSGSESLNEMITVDPDKVNVYAKRKREVGQAAVEWQALLSEAVL		
	RGQAVLVANSSPPPEPLDLMHKATISGIRSTITTLRALGADPAISLPDAASAPLRTIT		
	ADVFCKLPRVYSNPLRGLKLYTGEACRGDR"		
	217..288		
	/note="erythropoietin signal peptide"		
	289..792		
	/product="erythropoietin"		
BASE COUNT	297 a 455 c 433 g 277 t		
ORIGIN	unreported.		
Alignment Scores:			
Pred. No.:	5.26e-80	Length:	1462
Score:	899.50	Matches:	176
Percent Similarity:	95.34%	Conservative:	8
Best Local Similarity:	91.19%	Mismatches:	8
Query Match:	90.13%	Indels:	1
Db:	9	Gaps:	1
QY	1	MetGlyValHisGluCysProAlaIrrPheLeuThrPheLeuLeuSerLeuLeuSerLeuPro	20
Db	217	ATGGGGGTGACAGATGCTCCCTCGGCTGTGGGCTTCTCTCTGTCGTGCTCCCT	276

DB	Accession	Definition	Length	Score	Best Local Similarity	Query Match
OY	21	LeuGlyLeuProValLeuGlyAlaProPheGlyLeuLeuCysAspSerArgValLeuGlu	40	899.50	95.34%	8
Db	277	CHGGGCTCCAGTCCCGGGCCGCCACACGCCCTCATCTGTGACAGCGGATCTCTGGAG	336	90.13%	90.13%	1
OY	41	ArgTyrLeuLeuGluAlaLysGluAlaGluAsnIleThrThrGlyCysAlaGluHisCys	60	899.50	95.34%	8
Db	337	AGGTACTCTTGTGAAGGCCNAGGAGGCCCGAGATGTACAGATGGCTGTTCGAAAGCTGC	396	90.13%	90.13%	1
OY	61	SerLeuAsnGluAsnIleThrValProAspThrLysValAsnPheTyrAlaTrpLysArg	80	899.50	95.34%	8
Db	397	ACCTTGAAATGAGATATACACCTCCAGACACCAAGTTAACTTCTATCTCCGGAAGAGG	456	90.13%	90.13%	1
OY	81	MetGluValArgGluGlnAlaValGluValTrpGlnGlyLeuAlaLeuDeuSerGluAla	100	899.50	95.34%	8
Db	457	ATGGAGAGTCCGGGACAGAGGCTGTAGAGCTGTGGCAGGGGCTGGCTGTCTCAGAACTC	516	90.13%	90.13%	1
OY	101	ValLeuArgGlyGlnAlaLeuLeuValAsnSerSerGlnProTrpGluProLeuGlnLeu	120	899.50	95.34%	8
Db	517	GTCTCGCGGGGCGGAGGCCCTGTGGCCAACTCTTCCACAGCTTTCGAGGCCCTGCAGCTG	576	90.13%	90.13%	1
OY	121	HisValAspLysAlaValSerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGly	140	899.50	95.34%	8
Db	577	CACATGATAAAGCCATCATGTGGCCCTTCGACATCATCCACTCTTCTGGGCGCTGGGA	636	90.13%	90.13%	1
OY	141	AlaGlnLysGluAlaIleSerProPheAspAlaAlaSerAlaAlaProLeuArgThrIle	160	899.50	95.34%	8
Db	637	GGCCAG--GAAGCATCTCCCTCCCAATGGGGCTGGCTGCTCCATCCGAACATC	693	90.13%	90.13%	1
OY	161	ThrAlaAspThrPheArgLysLeuPheArgValTyrSerAsnPheLeuArgGlyLysLeu	180	899.50	95.34%	8
Db	694	ACTGCTGACACTTCTTGCAACTCTTCGAGCTCACTCAATTTCTCCGGGAAAGCTG	753	90.13%	90.13%	1
OY	181	LysLeuTyrThrGlyGluAlaCysArgThrGlyAspArg	193	899.50	95.34%	8
Db	754	AAGCTGTACACGGGAGGCGCTGCAGAGAGGGGACAGA	792	90.13%	90.13%	1

DB	Accession	Gene	Protein	Length	Score	Similarity	Local Similarity	Match	DB
Db	1770	CTGGGCTCTCCAGTCCCGGGGCGCCACACAGCCCTCATCTGTGACAGCCGATCCTGGAG		1829					
QY	41	ArgTYrLeuLeuGluAlaIleGlyGluIaGluAsnIleThrThrGlyCysAlaGluHisCys		60					
Db	1830	AGGTACCTCTTGGAGGCCAAGAGGAGCCAGATGTCACGATGGCGCTGTCCCAAGACTGC		1889					
QY	61	SerLeuAsnGluAsnIleThrValProAspThrLysValAspPheThrAlaTrpLysArg		80					
Db	1890	AGCTTGAATGAGAAATATACCTCCACAGCACCAAGTTAACTTGTATGCTCGAAGAGG		1949					
QY	81	MetGluValArgGlnGlnAlaValGluValTrpGlnGlyLeuAlaLeuLeuSerGluAla		100					
Db	1950	ATGGAGGTCCTGGGACAGGCTGTATAAGTCTGGAGGGGCGCTGGCCCTGCTCACAAGCT		2009					
QY	101	ValLeuAlaGlyGlnAlaLeuLeuValAsnSerSerGlnProTrpGluProLeuGlnLeu		120					
Db	2010	GTCCGCGGGGCGGCAAGCGCGTGTGGCCAACTCTTCCAGCCCTTTGAGGCCCTCGACG		2069					
QY	121	HisValAspLysAlaValSerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGly		140					
Db	2070	CACATGGATTAAGGCCATCATAGTGGCTTGCACACATCACCATCTCTTGGGCGCTGGGA		2129					
QY	141	AlaGlnLysGluAlaIleSerProProAspAlaIleSerAlaAlaProLeuArgThrIle		160					
Db	2130	GCCAG---GAAGCATCTCCCTCCACATGGGGGCTGGGCTGCTCCACTCGAAGCATC		2186					
QY	161	ThrAlaAspThrPheArgLysLeuPheArgValTyrSerAsnPheLeuArgGlyLysLeu		180					
Db	2187	ACTGCTGACACTTCTGCGAAMACTTTCGAGTCTCAATTTCTCTCGGGGGAAGCTG		2246					
QY	181	LysLeuTyrThrGlyGluAlaCysArgThrGlyAspArg		193					
Db	2247	AAGCTGTACACGGGGGAGGCTGTCCAGAGAGGGGACAGA		2285					
RESULT 14									
LOCUS	ARI58347		5107 bp	DNA		linear		PAT 17-OCT-2001	
DEFINITION	Sequence 3 from patent US 6251433.								
ACCESSION	ARI58347								
VERSION	ARI58347.1								
KEYWORDS									
SOURCE	Unknown.								
ORGANISM	Unknown.								
REFERENCE	1 (bases 1 to 5107)								
AUTHORS	Zuckermann,R.N., Dubois-Stringfellow,N., Dworki,V., Innis,M.A., Murphy,J.E., Cohen,F.E. and Uno,T.								
TITLE	Polycationic polymers								
JOURNAL	Patent: US 6251433-A 3 26-JUN-2001;								
FEATURES	Location/Qualifiers								
source	1..5107								
	/organism="unknown"								
BASE COUNT	1145 a 1463 c 1335 g 1164 t								
ORIGIN									
Alignment Scores:									
Pred. No.:	2.25e-79								
Score:	899.50								
Percent Similarity:	95.34%								
Best Local Similarity:	91.19%								
Query Match:	90.13%								
DB:	6								
	Gaps:								
	1								

US-09-813-775C-2 (1-193) x A92667 (1-5107)

Oy 1 MetGlyValHisGlyCysProAlaTrpLeuTrpLeuLeuSerLeuLeuSerLeuPro 20
Db 1710 ATGGGGGCGACAGAAATGCTCTGCTGCTGCTGCTTCTCTCTGCTGCTGCTCT 1766
Oy 21 LeuGlyLeuProValLeuGlyAlaProProArgLeuIleCysAspSerArgValLeuGlu 40

6y 1 MetGlyValHisGluCysProAlaTrpLeuTrrPheLeuLeuSerLeuLeuSerLeuPro 20
Db 1710 ATGGGGGTGCACCAATCTCTGGCTGGCTGTCTCTCTCTCGTGTCGCTCCCT 1765

Oy 21 LeuGlyLeuProValLeuGlyAlaProProArgLeuILecysAspSerArgValLeuGlu 40
Db 1770 CTGGAGCTCCCAAGTCCGGGGGCCACCACAGGCTCATTTGGACAGCGAGACTCTGGAG 1825

```

QY 41 ArgTyrLeuLeuGluAlaIylsGluAlaGlnIleThrThrGlyCysAlaGlnHisCys 60
Db 1830 AGGTACTCTTGGAGGCGCAAGAGAGCGGAGATGTACAGATGGCTGTCTCCGAAGCTGC 1889
QY 61 SerLeuasnGluasnIleThrValProAspThrLysValasnPheTyrAlaTrpLysArg 80
Db 1890 AGCTTGATGATGATATATACCGCTCCAGACACCAAGTTAACTTCTATGCTGGAGAGG 1949
QY 81 MetGluValArgGlnGlnAlaValaGluValTrpGlnGlyLeuAlaLeuSerGluAla 100
Db 1950 ATGGAGGTCCGGGACGAGCGGTGTAGAAAGTCTGGCAGGGGCTGGGCTGTCTCGAAAGCT 2009
QY 101 ValLeuArgGlyGlnAlaLeuValasnSerSerGlnProTrpGluProLeuGlnLeu 120
Db 2010 GTCTCGGGGGCCAGCGCGTGTGGCCACACTTCCAGCGCTTCGAGCCCTGCACACTG 2069
QY 121 HisValAspLysAlaValaSerGlyLeuArgSerLeuThrThrLeuAlaValaLeuGly 140
Db 2070 CACATGGATTAAGCATCATGTGGCTTCGAGCATCACACTGTCTGTGGGCGCTGGGA 2129
QY 141 AlaGlnLysGluAlaIleSerProAspAlaIleSerAlaIleProLeuArgThrIle 160
Db 2130 GCCCAG---GAAGCATCTCCCTCCGAGATCGGGCTCGCTGCTCCACTCCGAACCATC 2186
QY 161 ThrAlaAspThrPheArgLysLeuPheArgValTyrSerAsnPhelLeuArgGlyLysLeu 180
Db 2187 ACTGCTGACACTTCTTCGCAAACTCTTCGAGTCTACTCCAAATTCTCTCCGGGGAAGCTG 2246
QY 181 LysLeuTyrThrGlyGluAlaCysArgThrGlyAspArg 193
Db 2247 AAGCTGTACACGGGGAGGCGCTGCAGGAGAGGGGACAGA 2285

RESULT 15
LOCUS MACERYTHRO 691 bp mRNA linear PRI 14-OCT-1993
DEFINITION Macaca mulatta erythropoietin mRNA, complete cds.
ACCESSION U10609
VERSION 1.0
KEYWORDS Macaca mulatta kidney cDNA to mRNA.
SOURCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
ORGANISM Macaca mulatta
REFERENCE
AUTHORS Men,D., Boissel,J.P., Tracy,T.E., Gruninger,R.H., Mulcahy,L.S.,
Czelusniak,J., Goodman,M. and Bunn,H.F.
TITLE Erythropoietin structure-function relationships: high degree of
sequence homology among mammals
JOURNAL Blood 82 (5), 1507-1516 (1993)
MEDLINE 93372347
PUBMED 8364201

FEATURES
source Location/Qualifiers
1..691
/organism="Macaca mulatta"
/db_xref="taxon:9544"
/tissue_type="kidney"
1..579
/feature="precursor"
/codon_start=1
/product="erythropoietin"
/protein_id="AAA36842.1"
/db_xref="GI:342096"
/translation="MGVHECPAALMLLSLVSPLGLPVGAPRLVCDRLVRYLL
EAKAEVNTMGSESCSLMENTITPDTKYTFYAKRIEYGOQAVEYMOGALLSEAYL
RGQAVLANSQSPFEPQLMDKAISSIRISITITLLRALGQDAEATISLPDASAPLRTIT
ADTFCKLFRVYSNFKKLLKLTITGEACRRGDR"
1..81
82..576
/product="erythropoietin"
/evidence="experimental"
580..>691
sig_peptide
mat_peptide
3'UTR

```

```

BASE COUNT 134 a 234 c 189 g 134 t
ORIGIN
Alignment Scores:
Pred. No.: 6,92e-80 length: 691
Score: 894.50 Matches: 174
Percent Similarity: 95.34% Conservative: 10
Best Local Similarity: 90.16% Mismatches: 8
Query Match: 89.63% Indels: 1
DB: 9 Gaps: 1

US-09-813-775c-2 (1-193) x MACERYTHRO (1-691)
QY 1 MetGlyValHisGluCysProAlaTrpLeuThrLeuLeuSerLeuSerLeuPro 20
Db 1 ATGGGGGGGCGACGATATGCTCCTGGCTGGCTGCTTCCGTCTCTGCTCCCTCCCT 60
QY 21 LeuGlyLeuProValaLeuGlyAlaProProArgLeuIleCysAspSerArgValLeuGlu 40
Db 61 CTGGGCGCTCCCGATCCCGGGGGCGCCACACAGCGCTGTGTGACAGCGAGTCTGGAG 120
QY 41 ArgTyrLeuLeuGluAlaIylsGluAlaGlnIleThrThrGlyCysAlaGlnHisCys 60
Db 121 AGGTACTCTTGGAGGCGCAAGAGCGGAGATGTACAGATGGCTGTCTCCGAAGCTGC 180
QY 61 SerLeuasnGluasnIleThrValProAspThrLysValasnPheTyrAlaTrpLysArg 80
Db 181 AGCTTGATGATGATATATACCGCTCCAGACACCAAGTTAACTTCTATGCTGGAGAGG 240
QY 81 MetGluValArgGlnGlnAlaValaGluValTrpGlnGlyLeuAlaLeuSerGluAla 100
Db 241 ATGAGAGTCCGCGAGCGGCTGTAGAAAGTCTGGCAGGGGCTGTCTCTCGAAAGCT 300
QY 101 ValLeuArgGlyGlnAlaLeuValasnSerSerGlnProTrpGluProLeuGlnLeu 120
Db 301 GTCTCGGGGGCCAGCGCGTGTGGCCAACTTTCGAGCTTCGAGCCCTGCAGAGTG 360
QY 121 HisValAspLysAlaValaSerGlyLeuArgSerLeuThrThrLeuAlaValaLeuGly 140
Db 361 CACATGGATTAAGCATCATGTGGCTTCGAGCATCACACTGTCTGTGGGCGCTGGGA 420
QY 141 AlaGlnLysGluAlaIleSerProAspAlaIleSerAlaIleProLeuArgThrIle 160
Db 421 GCCCAG---GAAGCATCTCCCTCCGAGATCGGGCTGTCTCCACTCCGAACCATC 477
QY 161 ThrAlaAspThrPheArgLysLeuPheArgValTyrSerAsnPhelLeuArgGlyLysLeu 180
Db 478 ACTGCTGACACTTCTTCGCAAACTCTTCGAGTCTACTCCAAATTCTCTCCGGGGAAGCTG 537
QY 181 LysLeuTyrThrGlyGluAlaCysArgThrGlyAspArg 193
Db 538 AAGCTGTACACGGGGAGGCGCTGCAGGAGAGGGGACAGA 576

```

Search completed: January 8, 2003, 06:57:29
 Job time: 2607.65 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 8, 2003, 03:43:33 ; Search time 209.666 Seconds
(without alignments)
2072.993 Million cell updates/sec

Title: US-09-813-775c-2
Perfect score: 998
Sequence: 1 KGVHCPAMWLLLSLSLP.....NFLRGKLYTGACTGDR 193

Scoring table:
BL0SUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+pn.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09813775.r/unat.07012003.153120.23784/app.query.fasta.1.718
-DB=N.Geneseq.101002 -OFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bl0sum62 -TRANS=human0.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=ptc -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09813775.ecn1.1.0_4runat.07012003.153120.23784 -NCPU=6 -ICPU=3
-NO_XUPXY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEDOUT=120
-NARN_TIMEDOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
N.Geneseq.101002.*
1: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SID2/gcgdata4/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query length	DB ID	Description
1	998	100.0	585	22 AAC66882	Chimpanzee erythro
2	991	99.3	582	21 AAA46697	DNA encoding a hum
3	991	99.3	582	20 AAA46697	Human erythropoiet
4	991	99.3	616	20 AAX77361	Polynucleotide seq
5	991	99.3	616	21 AAA90393	DN encoding synthe
6	991	99.3	625	15 AAQ74760	Human prepro eryth
7	991	99.3	788	17 AAT31529	Human erythropoiet
8	991	99.3	788	21 AAT31529	Human erythropoiet
9	991	99.3	823	18 AAT64847	Human EPO long cDN
10	991	99.3	1015	13 AAQ24282	Epo:IL-3 short, re
11	991	99.3	1072	13 AAQ24285	Epo:IL-3 flex. rec
12	991	99.3	1158	20 AAX25701	Human erythropoiet
13	991	99.3	1245	8 AAN70360	Sequence encoding
14	991	99.3	1789	16 AAQ92296	Erythropoietin cDN
15	991	99.3	3100	24 AABN6161	Polynucleotide seq
16	991	99.3	10524	20 AAX77355	Nucleotide sequenc
17	991	99.3	10524	21 AAA90387	Plasmid pSINrep5 E
18	987	98.9	606	21 AAB07253	Human EPO cDNA seq
19	986	98.8	582	21 ABL50865	Modified erythropo
20	982	98.4	1342	7 AAN60513	cDNA of clone lamb
21	979.5	98.1	614	24 ABL59290	Nucleotide sequenc
22	977	97.9	579	22 AAT71883	Erythropoietin gen
23	977	97.9	629	24 ABL59289	Nucleotide sequenc
24	977	97.9	629	24 ABL59291	Nucleotide sequenc
25	977	97.9	6079	22 AAT71882	PE/C plasmid DNA.
26	971	97.3	1255	7 AAN60519	cDNA of clone lamb
27	968	97.0	724	9 AAN81554	EPO 140B encoding
28	943	94.5	1286	24 AAD24593	Human erythropoiet
29	941	94.3	1342	6 AAN50443	EPO clone lambda-H
30	899.5	90.1	1344	6 AAN50348	Monkey erythropoie
31	899.5	90.1	1344	19 AAV30955	Monkey erythropoie
32	899.5	90.1	5107	19 AAV21685	Vector plasmid CMV
33	871.5	87.3	681	19 AAO72942	Monkey Erythropoie
34	849	85.1	514	22 AAD06893	Human erythropoiet
35	846	84.8	994	13 AAQ24281	IL-3:Epo short, re
36	846	84.8	1239	18 AAT93879	DNA for oligopepti
37	844	84.6	570	21 ABL50864	Modified erythropo
38	843	84.5	498	21 ABL50878	Modified erythropo
39	843	84.5	501	19 AAV31031	Human erythropoiet
40	843	84.5	533	6 AAN50346	Human recombinant
41	843	84.5	533	19 AAV30963	ECERO gene used in
42	843	84.5	551	6 AAN50345	Human recombinant
43	843	84.5	551	19 AAV30970	SCEPO gene based o
44	843	84.5	1051	13 AAQ24284	IL-3:Epo flex. rec
45	838	84.0	498	21 AAA48373	Non-glycosylated e

ALIGNMENTS

RESULT 1
AAC66882
ID AAC66882 standard; cDNA; 585 BP.
XX
AC AAC66882;
XX
DT 27-MAR-2001 (first entry)
XX
DE Chimpanzee erythropoietin coding sequence SEQ ID NO: 3.
XX
KW Chimpanzee; erythropoietin; EPO; hybridisation probe; gene therapy;
XX mapping; therapeutic agent; ss.
XX
OS Pan sp.
XX
PN WO200068376-A1.
XX
PD 16-NOV-2000.

XX 05-MAY-2000; 2000WO-US12370.
 PF AAA46697
 XX ID AAA46697 standard; DNA; 582 BP.
 PR 07-MAY-1999; 99US-0307307.
 XX 28-MAR-2000; 2000US-0307307.
 XX (GETH) GENENTECH INC.
 PA
 XX Desauvage F, Henner DJ;
 PI
 XX WPI: 2001-007393/01.
 DR
 XX
 XX Nucleic acids encoding chimpanzee erythropoietin, useful for treatment
 PT of e.g. anemia, also derived proteins, antibodies and modulators -
 CC
 CC The present invention provides the coding and protein sequences of
 CC chimpanzee erythropoietin (EPO). These sequences can be used in gene
 CC therapy, to block the activity of EPO, as hybridisation probes, in
 CC genetic and chromosome mapping and as therapeutic agents.
 XX

Sequence 585 BP; 115 A; 185 C; 168 G; 117 T; 0 other;

Alignment Scores:
 Pred. No.: 3.6e-99 Length: 585
 Score: 998.00 Matches: 193
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-09-813-775C-2 (1-193) x AAC66882 (1-585)

OY 1 MetGlyValHisGluCysProAlaTrpLeuTrpLeuLeuSerLeuLeuSerLeuPro 20
 DB 1 ATGGGGGTGCAGAAATGCTCTGCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 60
 OY 21 LeuGlyLeuProValLeuGlyAlaProProArgLeuIleCysAspSerArgValLeuGlu 40
 DB 61 CTGGGCTCTCCAGTCTGCTGGGCGCCACACCGCTCATCTGTGACAGCGAGTCTCTGGAG 120
 OY 41 ArgTrpLeuLeuGluAlaGlyAlaGluValAsnIleThrThrGlyCysAlaGluHisCys 60
 DB 121 AGGTACTCTTGGAGCCAGAGAGCCGAGAAATATCAGACGCGGCTGTCCGAAACCTGC 180
 OY 61 SerLeuAsnGluAsnIleThrValProAspThrLysValAsnPheTyrAlaTrpLysArg 80
 DB 181 ACCTGAATGAGAAATATCTACTGCTCCAGACCAAAAGTTAATTTCTATGCTCGAAGAGG 240
 OY 81 MetGluValArgGlnGlnAlaValGluValITrpgInGlyLeuAlaLeuLeuSerGluAla 100
 DB 241 AAGGAGGTCAGGACAGAGCCCTAGAAAGTCTGCGAGGGCTGCGCTCTCGAAGCT 300
 OY 101 ValLeuArgGlyGlnAlaLeuLeuValAsnSerSerGlnProTrpGluProLeuGlnLeu 120
 DB 301 GTCTCTGCGGGGCGAGCCCTGTGCTCAACTCTTCCACGCCGTGGAGGCCCTGCAAGCTG 360
 OY 121 HisValAspLysAlaValSerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGly 140
 DB 361 CATGTGGAATAAGCCGCTGAGCTGTGCGACCTCAGACACACTGCTTGGGCTCGGGA 420
 OY 141 AlaGlnLysGluAlaIleSerProProAspAlaIleSerAlaAlaProLeuArgThrIle 160
 DB 421 GCCCAAGAGGAGCATCTCCCTCCAGATGCGGCTCAGCTGCATCCGGAACAATC 480
 OY 161 ThrAlaAspThrPheArgLysLeuPheArgValTyrSerAsnPheLeuArgGlyLysLeu 180
 DB 481 ACTGCTGACACTTCCGAAACTCTTCCGAGTCACTCCAAATTTCTCTCCGGAAGACTG 540
 OY 181 LysLeuTyrTrhGlyValAlaCysArgThrGlyAspArg 193
 DB 541 AAGCTGTACACAGGGAGGCTGCGAGGACAGGGGACAGA 579

RESULT 2

AAAA6697

ID AAA46697 standard; DNA; 582 BP.

XX AAA46697;

DT 25-SEP-2000 (first entry)

DE DNA encoding a human erythropoietin polypeptide.

KW Human; erythropoietin; EPO; inhibitor; nuclear factor-kappaB; NF-kappaB;

KW multi-drug resistance gene; malignant hemopathy; solid tumour;

KW malignant blood disease; leukaemia; lymphoma; solid cancer; ds.

XX Homo sapiens.

FT Key Location/Qualifiers

FT CDS 1..582

FT /tag= a

FT /product= "erythropoietin"

PN WO200030587-A2.

PD 02-JUN-2000.

PF 24-NOV-1999; 99WO-FR02897.

PR 25-NOV-1998; 98FR-0014858.

XX (CNRS) CENT NAT RECH SCI.

XX Hirsch F, Haeflner A;

XX WPI: 2000-399901/34.

DR P-PSDB; AAY93638.

XX Treatment of haematological or solid tumours using an inhibitor of the

PT activation of nuclear factor-kappaB, particularly to prevent

PT development of resistance to chemotherapeutics

XX Claim 5; Page 28-29; 30pp; French.

PS The present sequence encodes a human erythropoietin (EPO) polypeptide.

CC The human growth hormone protein is used as an inhibitor of the

CC activation of nuclear factor-kappaB (NF-kappaB). The inhibitor inhibits

CC resistance gene (which contains binding sites for NF-kappaB within its

CC regulatory regions). The inhibitors are used to produce pharmaceuticals

CC which may be used in the treatment of malignant hemopathy or solid

CC tumours. The inhibitors are especially used to treat malignant blood

CC diseases (leukaemia, lymphoma) and solid cancers (of breast or ovary).

XX Sequence 582 BP; 113 A; 181 C; 170 G; 118 T; 0 other;

Alignment Scores:

Pred. No.: 2.06e-98 Length: 582
 Score: 991.00 Matches: 192
 Percent Similarity: 99.48% Conservative: 0
 Best Local Similarity: 99.48% Mismatches: 1
 Query Match: 99.30% Indels: 0
 DB: 21 Gaps: 0

US-09-813-775C-2 (1-193) x AAA46697 (1-582)

OY 1 MetGlyValHisGluCysProAlaTrpLeuTrpLeuLeuSerLeuLeuSerLeuPro 20
 DB 1 ATGGGGGTGCAGAAATGCTCTGCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 60
 OY 21 LeuGlyLeuProValLeuGlyAlaProProArgLeuIleCysAspSerArgValLeuGlu 40
 DB 61 CTGGGCTCTCCAGTCTGCTGGGCGCCACACCGCTCATCTGTGACAGCGAGTCTCTGGAG 120

QY 41 ArgTyrlLeuEngluAlaIalysGluAlaGluAsnIleThrThrGlyCysAlaGluHisCys 60
 DB 121 AGGTACCTCTTGGAGGCCAAGAGGCCCGACAATATCAGAGGGGCTGTGCTGAACACATGC 180
 QY 61 SerLeuAsnGluAsnIleThrValProAspThrLysValAsnPhetYrAlaTrpLysArg 80
 DB 181 AGCTTGAAATGAAATATCAGTCTCCACACACCAAAGTTATTTCTATGCTGGAGAGG 240
 QY 81 MetGluValArgGlnGlnAlaIalysGluValIrrpGlnGlyLeuAlaLeuSerGluAla 100
 DB 241 ATGAGAGCGGGGACAGAGCGCGTAGAAGTGTGGCAGGGGCTGGGCGTGTGCGAAAGCT 300
 QY 101 ValLeuArgGlyGlnAlaLeuValAsnSerSerGlnProTrrpGluProLeuGlnLeu 120
 DB 301 GTCTGTGGGGGCCAGCGCCCTGTGTCAACTCTCCAGCGCGGTGGAGCCCTCCAGCTG 360
 QY 121 HisValAspLysAlaIalysSerGlyLeuArgSerLeuThrThrLeuArgAlaLeuGly 140
 DB 361 CATGTGATTAAGCGGTAGTGGCTTGCGAGCTCACACTGTGCTTGGGGCTGTGGGA 420
 QY 141 AlaGlnLysGluAlaIalysSerProProAspAlaIalysAlaIalProLeuArgThrIle 160
 DB 421 GCCCAGAAAGAGGACATCTCCCTCCAGATCGGGCTGAGCTGCCACCTCGAACAATC 480
 QY 161 ThrAlaAspThrPheArgGlyLysLeuPheArgValIrrSerAsnPhelLeuArgGlyLysLeu 180
 DB 481 ACTGCTGACACTTTCGCCCAAACTCTTCGAGTCACTCAATTCCTCCCGGGAAAGCTG 540
 QY 181 LysLeuTyrrThrGlyGluAlaCysArgThrGlyAspArg 193
 DB 541 AAGCTGTACACAGGGAGGCTCGACAGACAGAGGACAGA 579
 RESULT 3
 ID AAH46972 standard; cDNA; 582 BP.
 AC AAH46972;
 DX 29-OCT-2001 (first entry)
 DE Human erythropoietin (EPO) cDNA sequence.
 XX
 KM Transgenic; pig; human; erythropoietin; EPO; milk; PMSG; hcg;
 KM chorionic gonadotrophic hormone; WAP promoter; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..582
 FT /*tag= a
 FT sig_peptide /*product= "EPO"
 FT /*tag= b
 FT mat_peptide 82..579
 FT /*tag= c
 XX
 PN WO200159074-A1.
 PD 16-AUG-2001.
 XX
 PF 28-JUN-2000; 2000MO-KR00675.
 XX
 PR 14-FEB-2000; 2000KR-0006888.
 PA (KORE-) REPUBLIC KOREA.
 PI Chang W, Park J, Seong H, Min K, Yang B, Im G, Lee Y, Lee C;
 PI Kim J;
 XX WPI; 2001-514656/56.
 DR P-PSDB; AAB85573.
 XX
 PT Producing transgenic porcine that secretes human erythropoietin (hEPO)

PT in milk, by introducing vector comprising hEPO genome into fertilized
 PT eggs of porcine to which PMSG and hcg were administered, and developing
 PT progeny -
 XX
 PS Claim 4; Fig 3; 21pp; English.
 CC
 CC The invention relates to producing transgenic pigs (P) that secrete
 CC human erythropoietin (hEPO) in milk. The method involves administering
 CC PMSG and human chorionic gonadotrophic hormone (hcg) into (P), collecting
 CC fertilized eggs after mating, injecting expression vector containing a
 CC 2.6 kb WAP promoter, hEPO genome and SV40 poly A DNA into male pronuclei,
 CC transplanting them in surrogate mother pig and allowing it to give birth.
 CC The method provides transgenic porcine capable of secreting hEPO in their
 CC milk, thus producing the expensive useful medicine at a low cost with
 CC stability on a large scale, giving a contribution to the improvement of
 CC human health. The present sequence represents a base sequence for a human
 CC EPO cDNA incorporated into the genome of porcine.
 CC
 XX
 SQ Sequence 582 BP; 113 A; 181 C; 170 G; 118 T; 0 other:
 Alignment Scores:
 Pred. No.: 2,06e-98 Length: 582
 Score: 991.00 Matches: 192
 Percent Similarity: 99.48% Conservative: 0
 Best Local Similarity: 99.48% Mismatches: 1
 Query Match: 99.30% Indels: 0
 DB: 22 Gaps: 0
 US-09-813-775c-2 (1-193) x AAH46972 (1-582)
 QY 1 MetGlyValHisGlyCysProAlaTrpLeuThrLeuLeuSerLeuLeuSerLeuPro 20
 DB 1 ATGGGGGTGCACGAATGTCCTGGCTGTGGCTTCCCTGCTCCGCTGCTGCTCCT 60
 QY 21 LeuGlyLeuProValLeuGlyAlaProProArgLeuIleCysAspSerArgValLeuGlu 40
 DB 61 CTGGGCTCCCACTGCTGGGGCCGCCACACGCTCATCTGTGACAGCGAGTCTGGAG 120
 QY 41 ArgTyrlLeuEngluAlaIalysGluAlaGluAsnIleThrThrGlyCysAlaGluHisCys 60
 DB 121 AGGTACCTCTTGGAGGCCAAGAGGCCCGACAATATCAGAGGGGCTGTGCTGAACACATGC 180
 QY 61 SerLeuAsnGluAsnIleThrValProAspThrLysValAsnPhetYrAlaTrpLysArg 80
 DB 181 AGCTTGAAATGAAATATCAGTCTCCACACACCAAAGTTATTTCTATGCTGGAGAGG 240
 QY 81 MetGluValArgGlnGlnAlaIalysSerProProAspAlaIalysAlaIalProLeuArgThrIle 100
 DB 241 ATGAGAGTCCGGGACAGCGCGTAGAAGTGTGGCAGGGGCTGGGCTGTGCGAAAGCT 300
 QY 101 ValLeuArgGlyGlnAlaLeuValAsnSerSerGlnProTrrpGluProLeuGlnLeu 120
 DB 301 GTCTGTGGGGGCCAGCGCCCTGTGTCAACTCTCCAGCGCGGTGGAGCCCTCCAGCTG 360
 QY 121 HisValAspLysAlaIalysSerGlyLeuArgSerLeuThrThrLeuArgAlaLeuGly 140
 DB 361 CATGTGATTAAGCGGTAGTGGCTTGCGAGCTCACACTGTGCTTGGGGCTGTGGGA 420
 QY 141 AlaGlnLysGluAlaIalysSerProProAspAlaIalysAlaIalProLeuArgThrIle 160
 DB 421 GCCCAGAAAGAGGACATCTCCCTCCAGATCGGGCTGAGCTGCCACCTCGAACAATC 480
 QY 161 ThrAlaAspThrPheArgGlyLysLeuPheArgValIrrSerAsnPhelLeuArgGlyLysLeu 180
 DB 481 ACTGCTGACACTTTCGCCCAAACTCTTCGAGTCACTCAATTCCTCCCGGGAAAGCTG 540
 QY 181 LysLeuTyrrThrGlyGluAlaCysArgThrGlyAspArg 193
 DB 541 AAGCTGTACACAGGGAGGCTCGACAGACAGGAGACAGA 579
 RESULT 4
 ID AAH77361
 AAH77361 standard; DNA; 616 BP.

XX AAX77361;
 AC 09-AUG-1999 (first entry)
 DT
 DE Polynucleotide sequence of synthetic erythropoietin.
 XX
 XX Nucleic acid identification; exogenous protein; gene sorting;
 KM growth factor; membrane receptor; Sindbis virus; erythropoietin; ss.
 XX
 OS Synthetic.
 PN WO925876-A1.
 XX
 PD 27-MAY-1999.
 XX
 PF 17-NOV-1998; 98WO-US24520.
 XX
 PR 17-NOV-1997; 97US-0972218.
 XX
 PA (CYTO-) CYTOS BIOTECHNOLOGY GMBH.
 XX
 PI Bailey JE, Koller D, Orberger GH, Renner WA;
 XX
 DR WPI, 1999-357620/30.
 XX
 PT Isolating genes encoding proteins with selected properties, useful
 PT for identifying therapeutic agents or targets
 XX
 PS Disclosure: Fig 15; 136pp; English.
 XX
 CC The invention relates to the identification of a recombinant nucleic acid
 CC encoding an exogenous protein having a selected property. The method
 CC comprises preparing a population of eukaryotic host cells, culturing the
 CC cells under suitable conditions and identifying cells that contain the
 CC recombinant nucleic acid. The method is used to sort genes according to
 CC the type of proteins they express, and also to identify new ligand/
 CC receptor interactions. Typical applications of the nucleic acid and the
 CC exogenous protein are in isolation of new growth factors, cytokines,
 CC membrane receptors, cytoplasmic, organelle or nuclear proteins, all of
 CC which may be useful as therapeutic agents or therapeutic targets, e.g.
 CC apoptosis-promoting or tumour suppressing proteins, regulators of cell
 CC proliferation or metabolic processes etc. The protein can also be used to
 CC screen for specific modulators. The nucleic acid can also be used as
 CC sources of therapeutic antisense or ribozyme sequences. The method allows
 CC the protein (rather than a partial DNA sequence) to be isolated and,
 CC since a wide range of cells can be used, they can be expressed with the
 CC correct glycosylation pattern.
 XX
 SQ Sequence 616 BP; 118 A; 191 C; 184 G; 123 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,23e-98 Length: 616
 Score: 991.00 Matches: 192
 Percent Similarity: 99.48% Conservative: 0
 Best Local Similarity: 99.48% Mismatches: 1
 Query Match: 99.30% Gaps: 0
 DB: 20 Indels: 0
 US-09-813-775c-2 (1-193) x AAX77361 (1-616)
 QY 1 MetGlyValHisGluCysProAlaTrpLeuTrpLeuLeuSerLeuLeuSerLeuPro 20
 DB 16 ATGGGGGTGACGAAATGCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 75
 QY 21 LeuGlyLeuProValLeuGluAlaProPheArgLeuIleCysAspSerArgValLeuGlu 40
 DB 76 CTGGGCTCCAGTCTGGGGCCGCCACACGCTATCTGTGACAGCCGAGTCTGGAG 135
 QY 41 ArgTyrLeuLeuGluAlaLysGluAlaGlnIleThrThrGlyCysAlaGlnHisCys 60
 DB 136 AGGTACCTCTTGGAGGCCAAGAGGCGGAGATATACAGACGCGCTGCTGACACTGC 195

QY 61 SerLeuAsnGluAsnIleThrValProAspThrLysValAsnPhetYrAlaTrpLysArg 80
 DB 196 AGCTTGAAATGAAATATATCTGCTCCAGACACCAAAAGTTAATTCTATGCTGAAAGAG 255
 QY 81 MetGluValArgGlnGlnAlaValAlaTrpGlnGlyLeuAlaLeuSerGluAla 100
 DB 256 ATGGAGGTGGGCGAGGAGCCCTGAGAACTGTCGACAGGCTGGCCCTGCTGGAGAGCT 315
 QY 101 ValLeuArgGlyGlnAlaLeuValAsnSerSerGlnProTrpGluProLeuGlnLeu 120
 DB 316 GTCTGCGGGGCGAGGCGCTGTTGTCACACTCTCCACCGCTGGAGGCCCTGACAGCTG 375
 QY 121 HisValAspLysAlaValSerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGly 140
 DB 376 CATGTGATAAAGCCGCTGAGGCTTCGACAGCTCACCACCTGCTGCTGGGCTGGGGA 435
 QY 141 AlaGlnLysGluAlaIleSerProPheAspAlaAlaSerAlaAlaProLeuArgThrIle 160
 DB 436 GCCAGAAAGAAAGCCATCTCCCTCCAGATGCGGCTCAGCTCCTCCTGCAACCAATC 495
 QY 161 ThrAlaAspThrPheArgLysLeuPheArgValTyrSerAsnPhetLeuArgGlyLysLeu 180
 DB 496 ACTGCTGACACTTTCGCAAACTCTCCGAGTCTACTCCAAATTCTCCTCGGGGAAAGCTG 555
 QY 181 LysLeuTyrThrGlyGluAlaCysArgThrGlyAspArg 193
 DB 556 AAGCTGTACACAGGGGAGGCTGACAGACGAGGAGACAGA 594
 RESULT 5
 ID AAA90393 standard; DNA; 616 BP.
 XX
 XX AAA90393;
 AC 10-JAN-2001 (first entry)
 XX
 DR
 DE DN encoding synthetic erythropoietin (EPO).
 XX
 KW Nucleic acid identification; exogenous protein; drug screening;
 KW recombinant expression; synthetic erythropoietin; EPO; da.
 XX
 OS Synthetic.
 XX
 PN JP2000189173-A.
 XX
 PD 11-JUL-2000.
 XX
 PF 23-AUG-1999; 99JP-0236220.
 XX
 PR 17-NOV-1998; 98US-0193707.
 XX
 PR 17-NOV-1998; 98WO-US24520.
 XX
 PA (CYTO-) CYTOS BIOTECHNOLOGY GMBH.
 XX
 PI WPI, 2000-551637/51.
 XX
 PS Disclosure: Fig 15; 56pp; Japanese.
 XX
 CC The invention relates to the identification of a recombinant nucleic acid
 CC encoding an exogenous protein having a selected property. The method
 CC comprises preparing populations of eukaryotic host cells, where each cell
 CC comprises an expression vector encoding a different exogenous protein.
 CC The host cells are cultured under suitable conditions and the nucleic
 CC acid which encodes the exogenous protein is identified. The method is
 CC useful for the identification and isolation of proteins with a selected
 CC property. Typical applications of the nucleic acid and the exogenous
 CC protein are in isolation of new growth factors, cytokines, membrane
 CC receptors, cytoplasmic, organelle or nuclear proteins, all of which may


```

KM intravenous catheter clearance; blood clot; ss
XX Homo sapiens.
OS
XX US985607-A.
XX 16-NOV-1999.
XX PD
XX 27-JUN-1997; 97US-0883795.
XX PF
XX 19-DEC-1994; 94US-0358918.
XX PR
XX (CANG-) CANGENE CORP.
XX PA
XX Awang G, Delcuve G;
XX PI
XX WPI; 2000-012788/01.
XX DR
XX Recombinant DNA molecules encoding tissue plasminogen activator
PT proteins, operatively linked to a scaffold attachment region, useful
PT for the production of tissue plasminogen activator both in vivo and in
PT vitro -
PS
XX Example 2; Fig 4; 49pp; English.
XX
CC This sequence represents the human EPO long cDNA sequence.
CC The invention relates to a recombinant DNA molecule adapted for
CC expression of tissue plasminogen activator (tPA). The DNA molecule
CC comprise a sequence encoding tPA, an expression control sequence
CC operatively linked to the tPA sequence, and at least one human
CC apolipoprotein B scaffold attachment region (SAR) element (the SAR is not
CC a 5' proximal apolipoprotein B SAR). The SAR element is used to increase
CC the expression of the coding sequences. The recombinant nucleic acids may
CC be used for the recombinant production of tPA both in vitro or in vivo
CC (e.g. as part of a gene therapy procedure). TPA may be administered to
CC treat and remove blood clots. It is especially useful for the lysis of
CC occlusive coronary artery thrombi associated with evolving transmural
CC myocardial infarction to improve ventricular function and reduce the risk
CC of congestive heart failure. Additionally, it may be used in the
CC management of acute massive pulmonary embolism, Venous thrombosis and
CC acute ischemic stroke. Finally, tPA may be used in treating arterial
CC thrombosis or embolism, arteriovenous cannulae occlusion and intravenous
CC catheter clearance. In contrast to other plasminogen activators
CC (e.g. urokinase and streptokinase), the activity of tPA is relatively
CC localised and (in theory) is less likely to produce systemic haemorrhagic
CC disorders.
XX
SQ Sequence 788 BP; 145 A; 264 C; 228 G; 151 T; 0 other:

Alignment Scores:
Pred. No.: 3,12e-98 Length: 788
Score: 991.00 Matches: 192
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatch: 1
Query Match: 99.30% Indels: 0
DB: 21 Gaps: 0

US-09-813-775C-2 (1-193) x AAZ31999 (1-788)
QY 1 MetGlyValHisGluCysProAlaTrpLeuLeuSerLeuLeuSerLeuPro 20
Db 80 ATGGGGGTGCACGAATGTCTCGTGGCTTGTCCTCTCTGCTGCTGCTCTCT 139
QY 21 LeuGlyLeuProValLeuGluLysAlaProPheArgLeuLeuCysAspSerArgValLeuGlu 40
Db 140 CTGGGGCTCCCAAGTCTCTGGGGCCGCCACCACAGGCTCATCTGTACACGCCAAGTCTTGAG 199
QY 41 ArgTyLeuLeuGluLysGluAlaGlnAsnIleThrThrGlycSalGlnHisCys 60
Db 200 AGCTACCTTTTGAGCCCAAGAGGCCCAAGAATATCACAGCGGCTGTGCTGAACAATTCC 259
QY 61 SerLeuAsnGluAsnIleThrValProAspThrLysValAsnDheTyralaTrpLysArg 80

```

Db	260	AGCTTGAAATGGAATATATCACTGTGCTCCAGACACCAAAAGTATATTCTTATGCTGTGAAGAGG	319
Oy	81	MetGluValArgGlnGlnAlaValGluValTyrpGlnGlyLeuAlaLeuLeuSerGluAla	100
Db	320	ATGCAGGTCGGGACAGCAGCGCGTGAAGTCTGGCAGGAGCCCGCCGCTGTGCGAAGCT	379
Oy	101	ValLeuAArgGlyGlnAlaLeuLeuValAsnSerSerGlnProTrrpGluProLeuGlnLeu	120
Db	380	GTCCGCGCGGGCCAGAGCCCTTGTGGTCAACCTTCCAGCCGCTGGAGCCCCCTGCACCTG	439
Oy	121	HisValAspLysAlaValSerGlyLeuLeuSerLeuThrThrLeuLeuArgAlaLeuGly	140
Db	440	CATGTGGATTAAAGCCGCTAGTGGCTTCGCAGCCTCACCACTCTGCTTGGGCTTGGA	499
Oy	141	AlaGlnLysGluAlaIleSerProProAspAlaAlaSerAlaAlaProLeuArgThrIle	160
Db	500	GCCGGAAGGAAGCAAGCATCTCCCTCCAGATGCGGGCCTCAGCTGCTCCACTGCCAATC	559
Oy	161	ThrAlaAspThrPheArgLysLeuPheArgValTyrSerAsnPheLeuAArgGlyLysLeu	180
Db	560	ACTGCTGCACACTTCCGCAAACTCTCCGACTCTACATTCATTTCTCCGCGGAAACCTG	619
Oy	181	LysLeuTyrThrGlnGlyAlaCysArgThrGlyAspArg	193
Db	620	AAGCTGTACACAGGGAGGCGCTGCAGACAGGAGGACAGA	658
RESULT 9			
AA64847	ID	AA64847	standard; cDNA; 823 BP.
XX	XX	AA64847;	
AC	XX	04-FEB-1998	(first entry)
DT	XX	Human erythropoietin cDNA.	
DE	XX		
KM	XX	Adeno associated virus; virion; vector: erythropoietin; muscle;	
KM	XX	anaemia; renal failure; cancer; chemotherapy; antiviral therapy;	
KW	XX	haemoglobinopathy; sickle cell disease; thalassaemia; gene therapy;	
ss.	XX		
OS	XX	Homo sapiens.	
XX	XX		
FH	XX	Key	Location/Qualifiers
FT	XX	CDS	52..633
FT	XX		/*tag= a
XX	XX	WO926336-A1.	
PN	XX	24-JUL-1997.	
PD	XX	17-JAN-1997;	97WO-US00980.
PF	XX	16-JAN-1997;	97US-0785750.
PR	XX	18-JAN-1996;	96US-0588355.
XX	XX	(AVIG-) AVIGEN INC.	
PA	XX	Kurtzman GJ, Podsakoff GW;	
PI	XX	WPI; 1997-385339/35.	
DR	XX		
XX	XX		
PT	XX	Adeno associated virions for expressing erythropoietin in muscle	
PT	XX	cells - useful to treat anaemia associated with chronic renal	
PT	XX	failure, cancer, cancer chemotherapy or antiviral therapy, or	
PT	XX	haemoglobinopathy	
PS	XX	Disclosure: Fig 1; 81pp; English.	
XX	XX		
CC	CC	The cDNA sequence for human erythropoietin (EPO) is given in	
CC	CC	AA64847. A novel claimed composition includes recombinant adeno-	
CC	CC	associated virus (AAV) virions comprising an AAV vector containing	
CC	CC	a nucleic acid molecule encoding EPO operably linked to control	


```

XX AAX25701:
AC
XX
XX 21-MAY-1999 (first entry)
DT
XX
XX Human erythropoietin homodimer fusion gene.
DE
XX
XX Human; erythropoietin dimer; trimer; polymer; fusion protein; cancer;
KW biological activity; anaemia; proliferation; differentiation; ss;
KW progenitor; leucocyte; granulocyte; blood; myelosuppressed patient.
XX
XX Homo sapiens.
OS
XX Synthetic.
PN MO9902710-A1.
XX
XX 21-JAN-1999.
PD
XX
XX 09-JUL-1998; 98WO-US13944.
XX
XX 03-FEB-1998; 98US-0018138.
PR 10-JUL-1997; 97US-0890929.
XX
XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
PA
XX
XX Sytkowski AJ;
PI
XX WPI; 1999-120911/10.
DR P-PDB; AAW99360.
XX
XX New fusion protein with increased activity comprising at least two
PT protein molecules - used to, e.g. treat erythropoietin related
PT deficiency states for treatment of anaemia
XX
XX Example 1; Fig 16A-C; 119pp; English.
XX
XX This sequence represents a human erythropoietin (EPO) homodimeric fusion
CC gene. The invention relates to the production of dimeric, trimeric or
CC polymeric fusion proteins with increased biological activity. The
CC fusion proteins are used to treat or prevent protein-related deficiency
CC states, specifically, where the protein is erythropoietin (EPO);
CC AAX25689), anaemia, but also for increasing proliferation,
CC differentiation and activity of haematopoietic progenitors (e.g.
CC increasing numbers of leucocytes and granulocytes in the blood of
CC myelosuppressed patients) or for treating cancer and other cell growth
CC disorders.
CC
CC Sequence 1158 BP; 228 A; 352 C; 351 G; 227 T; 0 other;
SQ
XX
XX Alignment Scores:
Pred. No.: 5.28e-98 Length: 1158
Score: 991.00 Matches: 192
Percent Similarity: 99.48% Conservative: 0
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 99.30% Indels: 0
DB: 20 Gaps: 0
XX
XX US-09-813-775C-2 (1-193) x AAX25701 (1-1158)
OY 1 MetGIValIHISGLUCysPProAlaTriPLeuTrPLeuLeuSerLeuLeuSerLeuPro 20
Db 11 ATGGGGGTCAGCAATGTCTGCCTGGCTGGCTTCTCTGTCCTCGCTGCTGCCCTCC 70
OY 21 LeuGIyLeuPProValIleuGIyAlAProPProArgLeuLIEcysASpSErAtyValIleGlu 40
Db 71 CTGGGCTCTCCCAAGTCTCTGGGGCCCCCACACAGCCTATCTGTACACAGCCAGTCTGGAG 130
OY 41 ArgTyrlLeuLeuGIyAlaLySGlualyGlualsInIleThrThrelycSalagluHiscys 60
Db 131 AGGTACTCTTGAGAGGCCAAAGCAGGCGGAGATATCATCACAGCGGGCTGTGAACACTGC 190
OY 61 SerLeuAsnGIyAsnIleThrValProAspThrItySAlasbnPhetyrralatriplysArg 80

```

Dd		191	AGCTGAATGAGAAATATCACTGTCCACAGCACCAAGAATTAAATTTCTATGGCTTGGAAAGG	250
Qy		81	MetGIuValArGIInGIAlAlaIcIuValITrPGInGIyLeuAlaLeuLauSerGIuAla	100
Dd		251	ANGGAGGTCGGGCACAGCGCCTAGAACGTGGCAGGCCCTGGCCCTGTCCGGAAGCT	310
Qy		101	VallAuArGgLyGIAlaLeuLeuValAuaAnSerSerGIInProTrPGIuProLeuGIInleu	120
Dd		311	GTCCTGCGGGGGCCAGGCCCTGTGGTCAACTTTCACCAGCCGTGGAGCCCTGCAGCTG	370
Qy		121	HISValAsPlySAlaIaValSerGIyLeuArGIserLeuThrTrhLeuLeuArGIAlaLeuGIy	140
Dd		371	CATGTGATAAAGCCGTCAGTGAGCCTTCGCACAGCTCACACACTCTGCTCGGGCTCGGGA	430
Qy		141	AlaGIuLySGIuAlaIIeSerProTroAspAlaIaSerAlaIaLaProLeuArGIThrIle	160
Dd		431	GCCCCAAGGAGAACCATCTCCCCTCCAATGGCGCTCAGCGCTCCACATCCGAACAATC	490
Qy		161	ThrAlaAsPTPhArGIyLeuLeuPheArGIyTyTrSerAsnPhaLeuArGIyLyseu	180
Dd		491	ACTGCTGACACTTTCGCAAACTCTTCGAGGTCTACTCCAAATTTCTCTCGGGGAAAAGCTG	550
Qy		181	LysLeuTYrThrGIyGIuAlaCyArGIthrGIyAsParG	193
Dd		551	AAGCTGTACACAGGGAGGCCCTGCAGACAGGGGACACAG	589
RESULT		13		
ID		AAAN70360		
XX		AAAN70360 standard; cDNA; 1245 BP.		
AC		AAAN70360;		
DT		19-FEB-1991 (first entry)		
DE		Sequence encoding human erythropoietin (Epo).		
XX		Renal anaemia therapy; hormone; ss.		
OS		Homo sapiens.		
FH		Key	Location/Qualifiers	
FT		sig_peptide	79..159	
FT		mat_peptide	/tag= a	
FT			160..660	
FT			/tag= b	
PN		EP232034-A.		
PD		12-AUG-1987.		
XX		19-JAN-1987;	87EP-0300399.	
XX		23-JAN-1986;	86JP-0012868.	
FA		(SUMO) SUMITOMO CHEM IND KK.		
PA		(SUMI-) SUMITOMI SEIYAKU KK.		
PI		Yanagi H, Ogawa I, Okamoto M, Hozumi T, Soga A, Yoshima T;		
PI		Tsutsumi M;		
DR		WPI: 1987-223006/32.		
DR		P-PSDB; AAP70256.		
PT		Human erythropoietin prodn. - by culturing human cells, esp.		
PT		Namawa cells, transformed with DNA encoding human erythropoietin		
XX		Disclosure: Fig 1; 22pp; English.		
CC		A cDNA library was prepd. from the poly (A) RNA, which was isolated		
CC		from the erythropoietin-producing human hepatoma cell Hp-1. The cDNA		
CC		library was screened using the probes given in AAAN70361 and AAAN70362. A		
CC		plasmid (named as pB8.420) was isolated. The nucleotide sequence of		
CC		the cDNA obtained from this clone is shown in AAAN70360.		

XX Sequence 1245 BP; 273 A; 384 C; 351 G; 237 T; 0 other;
 Alignment Scores:
 Pred. No.: 5.83e-98 Length: 1245
 Score: 991.00 Matches: 192
 Percent Similarity: 99.48% Conservative: 0
 Best Local Similarity: 99.48% Mismatches: 1
 Query Match: 99.30% Indels: 0
 Gaps: 0

US-09-813-775c-2 (1-193) x AAN70360 (1-1245)

QY 1 MetGlyValHisGluCysProAlaTTrpLeuTTrpLeuLeuSerLeuLeuSerLeuPro 20
 |||||||
 Db 79 ATGGGGGTGCACGAATGCTGCTGGCTGCTGCTTCTCTGCTGCTGCTGCTGCTGCT 138
 |||||||

QY 21 LeuGlyLeuProValLeuGlyAlaProProArgLeuIleCysAspSerArgValLeuGlu 40
 |||||||
 Db 139 CTGGCCCTCCAGTCTGGGGCCGCCACACGCTCATCTGTGACAGCGGAGTCTGGAG 198
 |||||||

QY 41 ArgTyrLeuLeuGluAlaValSerGluAlaGluAsnIleThrThrGlyCysAlaGluHisCys 60
 |||||||
 Db 199 AGGTAACTCTTGGAGGCCCAAGAGGCCGAGAAATATCACAGAGCGGCTGTGCTGAACACTGC 258
 |||||||

QY 61 SerLeuAsnGluAsnIleThrValProAspThrLysValAsnPheTyrAlaTTrpLysArg 80
 |||||||
 Db 259 AGCTTGAATGAGATATCATCTGCTCCAGACACCAAGTTATTTCTATGCTGGAAGAGG 318
 |||||||

QY 81 MetGlyValArgGlnGlnAlaValGluValTTrpGlnGlyLeuAlaLeuLeuSerGluAla 100
 |||||||
 Db 319 ATGAGAGTCTGGGCGACGAGCGCGTAGAAGTCTGGCAGGCGCTGCTGCTGCGAAACT 378
 |||||||

QY 101 ValLeuArgGlyGlnAlaLeuLeuValAsnSerSerGlnProTTrpGluProLeuGlnLeu 120
 |||||||
 Db 379 GTCTCGGGGGCGCCAGCGCTTGTGTCACTCTTCCAGCGCTGGAGCCCTGCAGCTG 438
 |||||||

QY 121 HisValAspLysAlaValSerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGly 140
 |||||||
 Db 439 CATGTGATTAAGCGCTGAGTGGCTTGCAGCTTGCACACTCTGCTTCCGGCTCTGGGA 498
 |||||||

QY 141 AlaGlnLysGluAlaIleSerProProAspAlaIleSerAlaIleProLeuArgThrIle 160
 |||||||
 Db 499 GCCCAGAAAGGACCATCTCCCTCCAGATGGCGCTCAGCTGCTCCACTCCGAACATTC 558
 |||||||

QY 161 ThrAlaAspThrPheArgLysLeuPheArgValTyrSerAsnPheLeuArgGlyLysLeu 180
 |||||||
 Db 559 ACTGCTACACTTCCGCAAACTCTTCCGAGTCTACTCCAACTTTCCTCCGGGAAAGCTG 618
 |||||||

QY 181 LysLeuTyrThrGlyGluAlaCysArgThrGlyAspArg 193
 |||||||
 Db 619 AAGCTGTACACAGGGAGGCTTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 657
 |||||||

RESULT 14
 AA0922296
 ID AA0922296 standard; cDNA; 1789 BP.
 XX
 AC AA0922296:
 XX
 DT 30-OCT-1995 (first entry)
 XX
 DE Erythropoietin cDNA.
 XX
 KM Erythropoietin; anemia; gene therapy; gene transfer; red blood cell;
 KW RBC; erythrocyte; transformation; myoblast; EPO; ss.
 XX
 OS Homo sapiens.
 XX
 FH key Location/Qualifiers
 FT CDS 625..1206
 FT /*tag= a
 XX
 PN M09513376-A.

XX
 PD 18-MAY-1995.
 XX
 XX
 PE 09-NOV-1994; 94MO-US13066.
 XX
 PR 07-OCT-1994; 94US-0320480.
 PR 10-NOV-1993; 93US-0149871.
 XX
 PA (AMGE-) AMGEN INC.
 PA (UYSC-) UNIV SOUTHERN CALIFORNIA.
 XX
 PI Hamamori Y, Kedes LH, Samal BB;
 DR WPI; 1995-194095/25.
 DR P-PSDB: AAR74141.
 XX
 PT Gene therapy for treatment of anaemia - and increasing red blood cell
 PT production by transforming red blood cells with the erythropoietin gene
 PS Disclosure; Page 38-40; 51pp; English.
 XX
 CC The cDNA sequence for human EPO is given in AA0922296. Delivery of
 CC the cDNA to target muscle cells, in vivo or in vitro, allows
 CC increased RBC prodn.
 XX
 SQ Sequence 1789 BP; 353 A; 637 C; 497 G; 302 T; 0 other;
 Alignment Scores:
 Pred. No.: 9.57e-98 Length: 1789
 Score: 991.00 Matches: 192
 Percent Similarity: 99.48% Conservative: 0
 Best Local Similarity: 99.48% Mismatches: 1
 Query Match: 99.30% Indels: 0
 Gaps: 0

US-09-813-775c-2 (1-193) x AA0922296 (1-1789)

QY 1 MetGlyValHisGluCysProAlaTTrpLeuTTrpLeuLeuSerLeuLeuSerLeuPro 20
 |||||||
 Db 625 ATGGGGGTGCACGAATGCTGCTGGCTGCTGCTTCTCTGCTGCTGCTGCTGCTGCT 684
 |||||||

QY 21 LeuGlyLeuProValLeuGlyAlaProProArgLeuIleCysAspSerArgValLeuGlu 40
 |||||||
 Db 685 CTGGGCCCTCCAGTCTGGGGCCGCCACACGCTCATCTGTGACAGCGGAGTCTGGAG 744
 |||||||

QY 41 ArgTyrLeuLeuGluAlaValSerGluAlaGluAsnIleThrThrGlyCysAlaGluHisCys 60
 |||||||
 Db 745 AGGTAACTCTTGGAGGCCCAAGAGGCCGAGAAATATCACAGAGCGGCTGTGCTGAACACTGC 804
 |||||||

QY 61 SerLeuAsnGluAsnIleThrValProAspThrLysValAsnPheTyrAlaTTrpLysArg 80
 |||||||
 Db 805 AGCTTGAATGAGATATCATCTGCTCCAGACACCAAGTTATTTCTATGCTGGAAGAGG 864
 |||||||

QY 81 MetGlyValArgGlnGlnAlaValGluValTTrpGlnGlyLeuAlaLeuLeuSerGluAla 100
 |||||||
 Db 865 ATGAGAGTCTGGGCGACGAGCGCGTAGAAGTCTGGCAGGCGCTGCTGCTGCGAAACT 924
 |||||||

QY 101 ValLeuArgGlyGlnAlaLeuLeuValAsnSerSerGlnProTTrpGluProLeuGlnLeu 120
 |||||||
 Db 925 GTCTCGGGGGCGCCAGCGCTTGTGTCACTCTTCCAGCGCTGGAGCCCTGCAGCTG 984
 |||||||

QY 121 HisValAspLysAlaValSerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGly 140
 |||||||
 Db 985 CATGTGATTAAGCGCTGAGTGGCTTGCAGCTTGCACACTCTGCTTCCGGGCTCTGGGA 1044
 |||||||

QY 141 AlaGlnLysGluAlaIleSerProProAspAlaIleSerAlaIleProLeuArgThrIle 160
 |||||||
 Db 1045 GCCCAGAAAGGACCATCTCCCTCCAGATGGCGCTCAGCTGCTCCACTCCGAACATC 1104
 |||||||

QY 161 ThrAlaAspThrPheArgLysLeuPheArgValTyrSerAsnPheLeuArgGlyLysLeu 180
 |||||||
 Db 1105 ACTGCTACACTTCCGCAAACTCTTCCGAGTCTACTCCAACTTTCCTCCGGGAAAGCTG 1164
 |||||||

QY 181 LysLeuTyrThrGlyAlaCysArgThrGlyAspArg 193
 DB 1165 AAGCTGTACACAGGGAGGCGCTGACAGACAGGGGACAGA 1203
 RESULT 15
 ABN86161
 ID ABN86161 standard; DNA; 3100 BP.
 AC ABN86161;
 XX
 DT 23-SEP-2002 (first entry)
 XX
 DE Nucleotide sequence of pP1666 plasmid.
 XX
 KW Expression system; molecular switch; steroid hormone receptor; transgene;
 KW erythropoietin; clotting factor; interferon; cyostatic; viroicide;
 KW cancer; gene therapy; ds.
 XX
 OS Synthetic.
 XX
 PN WO200224899-A2.
 XX
 PD 28-MAR-2002.
 XX
 PF 25-SEP-2001; 2001WO-US30305.
 XX
 PR 25-SEP-2000; 2000US-235030P.
 PR 10-JAN-2001; 2001US-260781P.
 PR 23-MAR-2001; 2001US-278281P.
 XX
 PA (VALE-) VALENTIS INC.
 XX
 PI Abruzzese RV, Mehta V, Nordstrom JL;
 DR WPI; 2002-537166/57.
 XX
 PT Novel inducible expression system for regulating nucleic acid
 PT expression in gene therapy, provides undetectable biological effect or
 PT gene expression in absence of inducer, and high expression in presence
 PT of inducer -
 XX
 PS Claim 43; Fig 14; 101pp; English.
 XX
 CC The invention provides an inducible expression system (a molecular switch
 CC inducible-expression system) that comprises first expression cassette
 CC (a molecular switch or GENESWITCH plasmid) having first nucleic acid (N1)
 CC encoding for FP (a molecular switch protein) comprising mutated DNA-
 CC binding domain, a transregulatory domain, and a mutated ligand-binding
 CC domain of a steroid hormone receptor; and second expression cassette
 CC (inducible gene plasmid) having second nucleic acid (N2) encoding desired
 CC product, where FP is activated in presence of inducer and expression of
 CC N2 is controlled by binding of activated FP on promoter region of N2. The
 CC expression system is useful for regulating expression of transgene in
 CC vivo. The system is most preferably useful for inducing the expression of
 CC N2 encoding erythropoietin, a clotting factor, or an interferon (IFN).
 CC IFN gene therapy is useful for treating cancer and chronic viral
 CC infections. The present sequence represents the complete nucleotide
 CC sequence of pP1666 plasmid.
 XX
 SO Sequence 3100 BP; 656 A; 913 C; 856 G; 675 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,03e-97 Length: 3100
 Score: 991.00 Matches: 192
 Percent Similarity: 99.48% Conservative: 0
 Best Local Similarity: 99.48% Mismatches: 1
 Query Match: 99.30% Indels: 0
 DB: 24 Gaps: 0
 US-09-813-775C-2 (1-193) x ABN86161 (1-3100)
 QY 1 MetGlyValHisGluCysProAlaTrpLeuTyrPheLeuLeuSerLeuLeuPro 20
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 417 ATGGGGGTGCACGAAATGCTCCCTGGCTGGCTGCTCTCTGCTCCCTGCTCCCT 476
 QY 21 LeuGlyLeuProValLeuGlyAlaProPheArgLeuIleCysAspSerArgValLeuGlu 40
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 477 CTGGGCTCCACAGTCTCTGGGGGGCCGACACAGCGCTCATCTGTACAGCCGGCTCGGAG 536
 QY 41 ArgTyrLeuLeuGlyAlaAlaYsglyAlaGluAsnIleThrThrGlyCysAlaGluHisCys 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 537 AGGTATCTCTGGAGGCGCAAGAGGCGCCGAGAAATATCAGCAGCGGCTGCTGAACTGC 596
 QY 61 SerLeuAsnGluAsnIleThrValProAspThrIleValAspPheTyrAlaTrpIleAspArg 80
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 597 AGCCTAAATGAGAAATATCACTGTGCCAGACCCAAAGTAAATTTCTATGCTGGAGAGG 656
 QY 81 MetGluValArgGlnGlnAlaValGluValTrrpGlnGlyLeuAlaLeuLeuSerGluAla 100
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 657 ATGGAGGTGGGGGAGCAGGCGCGTGAAGTCTGGCAGGGCTGGCCCTGCTGCCAAGCT 716
 QY 101 ValLeuArgGlyGlnAlaLeuLeuValAsnSerSerGlnProTrpGluProLeuGlnLeu 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 717 GTCCTCGGGGGCCAGGCGCTGCTGTCGCAACTCTTCCACGCCGCGGAGCCCTGCAGCTG 776
 QY 121 HisValAspIleValAlaValSerGlyLeuArgSerLeuThrThrLeuLeuArgAlaLeuGly 140
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 777 CATGTGATTAAGCCGTCGATGCGCGAGCTGCGACCTCACCACTCTGCTCGGCGCTGGGA 836
 QY 141 AlaGlnIleGlyAlaIleSerProPheAspAlaIleSerAlaAlaProLeuArgThrIle 160
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 837 GCCCAGAGGAAGCCATCTCCCTCCAGATGGGGCTCGCGTCTCCATCCGCAATC 896
 QY 161 ThrAlaAspThrPheArgIleLeuPheArgValTyrSerAsnPheLeuArgGlyLysLeu 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 897 ACTGCTGACACACTTCCGCAAACTCTTCCAGTCTACTCCAAATTCCTCCGGGAAAGCTG 956
 QY 181 LysLeuTyrThrGlyAlaCysArgThrGlyAspArg 193
 DB 957 AAGCTGTACACAGGGAGGCGCTGACAGACAGGGGACAGA 995
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Search completed: January 8, 2003, 05:26:57
 Job time : 222.666 secs

Gencore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 8, 2003, 05:16:34 ; Search time 1619.26 Seconds
(without alignments)
1930.339 Million cell updates/sec

Title: US-09-813-775C-2

Perfect score: 998

Sequence: 1 MGVECPAMIMLLLSLSP.....NPLRGKIKLYTGACRTGDR 193

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=Frame+ p2n.model -DEV=xih
-O=/cgrn2.1/USPTO.spool/US09813775/runat_07012003_153122_23837/app.query.fasta_1.718
-DB=EST -OPMT=fastap -SUFFIX=p2n.rst -MINMATCH=0 -ILOOPC=0 -ILOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pic -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US09813775 -CGCN_1_1_821 -runat_07012003_153122_23837 -NCPU=6 -ICPU=3
-NO_XIPY -NO_MMAP -LARGEJOURY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-MAIN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database :
EST:
1: em.estdb:*
2: em.esthum:*
3: em.estlin:*
4: em.estm:*
5: em.estov:*
6: em.estpl:*
7: em.estro:*
8: em.htc:*
9: gb.est1:*
10: gb.est2:*
11: gb.htc:*
12: gb.est3:*
13: gb.est4:*
14: gb.est5:*
15: em.estfun:*
16: em.estom:*
17: gb.gss:*
18: em.gss.hum:*
19: em.gss.inv:*
20: em.gss.pln:*
21: em.gss.vrt:*
22: em.gss.fun:*
23: em.gss.man:*
24: em.gss.mus:*
25: em.gss.other:*
26: em.gss.pro:*
27: em.gss.rod:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query	Match Length	DB ID	Description
1	443.5	44.4	567	9	AA662379
2	180	18.0	686	13	BM438685
3	159	15.9	444	10	BE663803
4	139	13.9	968	17	CNS02F50
5	117	11.7	365	13	BM279790
6	117	11.7	611	13	BM315886
7	99.5	10.0	967	12	BG475544
8	94.5	9.5	867	12	BF184120
9	94	9.4	803	13	BI648123
10	92	9.2	593	14	BM931504
11	92	9.2	657	14	BQ185758
12	91.5	9.2	800	13	BI909650
13	91	9.1	639	13	BJ528644
14	90	9.0	912	10	BE618292
15	88.5	8.9	609	9	AA538257
16	88.5	8.9	940	14	BQ684126
17	88.5	8.9	1035	17	CNS02175
18	87.5	8.8	955	13	BM006435
19	87.5	8.8	976	14	BQ642885
20	87.5	8.8	1494	11	AK008748
21	87	8.7	299	10	BE061329
22	87	8.7	814	9	AU130785
23	87	8.7	1149	13	BM048467
24	86.5	8.7	713	14	BM723615
25	86.5	8.7	770	14	BQ751539
26	86.5	8.7	809	14	BQ750964
27	86.5	8.7	1091	13	BM551902
28	86	8.6	535	14	BQ658306
29	86	8.6	606	14	BM947371
30	86	8.6	671	12	BG675678
31	86	8.6	854	13	BI600127
32	85.5	8.6	663	10	BB653289
33	85.5	8.6	1039	9	BI251034
34	85	8.5	458	9	AI030233
35	85	8.5	568	13	BI682360
36	85	8.5	641	13	BI594824
37	85	8.5	686	9	AI722105
38	84.5	8.5	644	13	BI601795
39	84.5	8.5	675	17	BH466321
40	84.5	8.5	721	17	AZ981594
41	84.5	8.5	749	12	BE916828
42	84.5	8.5	866	12	BG340812
43	84.5	8.5	874	12	BF674417
44	84.5	8.5	979	10	BB029809
45	84.5	8.5	1149	12	BG389982

ALIGNMENTS

RESULT 1
AA662379/c
LOCUS
DEFINITION nu93g01.s1 NCI-CGAP_P122 Homo sapiens cDNA clone IMAGE:121888 3'
similar to gb:S65458 ERYTHROPOIETIN PRECURSOR (HUMAN);, mRNA
ACCESSION
VERSION AA662379
KEYWORDS
SOURCE AA662379.1 GI:2616470
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 567)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
Emmert-buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbip/image/image.html
Seq primer: 40m13 fwd. ET from Amersham
High quality sequence stop: 375.
Location/Qualifiers
1. 567
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1ib="IMAGE:1218288"
/clone_1ib="NCI-CGAP_P122"
/sex="male"
/tissue_type="normal prostate"
/lab_host="DH10B"
/note="Organ: prostate; Vector: pT73D-Pac (Pharmacia)
with a modified polylinker; 1st strand cDNA was prepared
from normal prostate bulk tissue, and was then primed with
a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library is normalized, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 125 a 142 c 181 g 119 t

ORIGIN

Alignment Scores:
Pred. No.: 6,06e-41 Length: 567
Score: 443.50 Matches: 111
Percent Similarity: 89.608 Conservative: 1
Best Local Similarity: 88.804 Mismatches: 12
Query Match: 44.444 Indels: 5
DB: 9 Gaps: 1

US-09-813-775C-2 (1-193) x AA662379 (1-567)

QY 69 ProAspThrLysValAsnPheryrAlaTrpLysArgMetGluValAlaArgGlnAlaVal 88
|||||
DB 567 CCAGACACCAAGGTAA-TCCTATGCTCGCAAAAGTAAGT---CGGCGACGACCCGTA 512
QY 89 GluValITPcInGlyLeuAlaLeuSerGluAlaValLeuArgGlyGlnAlaLeu 108
|||||
DB 511 GAAGTCTGA--GGGCTGGCCCTTCTGTGCAAAAGTCTCTGCG--GGCCATGCCCTGTG 455
QY 109 ValAsnSerGlnProTrpGluProLeuGlnLeuHisValAspLysAlaValSerGly 128
|||||
DB 454 TCAAACTCTTCCAGCCGCGGAGGCCCTGACCTCATGTGATAAAGCCGTCAGTGGC 395
QY 129 LeuArgSerLeuThrThrLeuLeuArgAlaLeuGlyAlaGlnLysGlnAlaLeuSerPro 148
|||||
DB 394 TT--CGAGCCTCACCTCGCTGGGCTGCGGAGCCGAGAAAGGACGATCTCCCT 336
QY 149 ProAspAlaAlaSerAlaAlaProLeuArgThrIleThrAlaAspThrPheArgLysLeu 168
|||||
DB 335 CCGAATGGCGCCCTCAGCTGCTCCACATCGAATCATGCTGACACTTCCGCAAACTC 276
QY 169 PheArgValITrSerAsnPhelLeuArgGlyLysLeuLysLeuThrArgGlyAlaGly 188
|||||
DB 275 TTCGGAGTCTACCTCCAAATTTCTCCGCGGAAAGCTGAAGCTGTACACAGGAGGCTGC 216
QY 189 ArgThrGlyAspArg 193
|||||
DB 215 AGGACAGGCGACAGA 201

RESULT 2
BM438685
LOCUS
DEFINITION
IP1v01464 Liver cDNA library Ictalurus punctatus cDNA 5' similar
to Erythroprotein, mRNA sequence.
ACCESSION
BM438685
VERSION
BM438685.1 GI:18460407
KEYWORDS
EST.
ORGANISM
channel catfish.
Ictalurus punctatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
Ictaluridae; Ictalurus.
1 (bases 1 to 686)
REFERENCE
Feng, J., Kucuktas, H., Kocabas, A., Li, P. and Liu, Z.
Transcriptome of channel catfish (Ictalurus punctatus): initial
analysis of expressed sequence tags from the liver
Unpublished (2002)
JOURNAL
CONTACT: Liu ZJ
The Fish Molecular Genetics and Biotechnology Laboratory,
Department of Fisheries and Allied Aquacultures and Program of Cell
and Molecular Biosciences
Auburn University
203 Swingle Hall, Auburn University, Auburn, AL 36849, USA
Tel: 334 844 4054
Fax: 334 844 9208
Email: zliu@acesag.auburn.edu
Seq primer: M13 Reverse.
Location/Qualifiers
1. 686
/organism="Ictalurus punctatus"
/db_xref="taxon:7998"
/clone_1ib="Liver cDNA library"
/note="Organ: Liver; Vector: pSport1; Site_1: NotI;
Site_2: SalI"

BASE COUNT 189 a 171 c 155 g 171 t

ORIGIN

Alignment Scores:
Pred. No.: 2.01e-10 Length: 686
Score: 180.00 Matches: 45
Percent Similarity: 51.098 Conservative: 25
Best Local Similarity: 32.858 Mismatches: 47
Query Match: 18.044 Indels: 20
DB: 13 Gaps: 3

US-09-813-775C-2 (1-193) x BM438685 (1-686)

QY 56 CysAlaGluHisCysSerLeuAsnGluAsnIleThrValProAspThrLysValAsn 75
|||||
DB 15 TGTACAGAGGAATGAGATGATGCTGTACCTCTCAGTCTGACAGAGTGGATTTT 74
QY 76 TyrAlaITrPlysArgMetGluValArgGlnGlnAlaValGluValITrPcInGlyLeuAla 95
|||||
DB 75 CATGTTGGGAAAGCAAAATACAGAGAGCAGGCTCAAGAACTGCAAGACGACTATGG 134
QY 96 LeuLeuSerGlnAlaValLeuArgGlyGlnAlaLeuLeuValAsnSerGlnProTrp 115
|||||
DB 135 TTCCTAAACATGGCCATTCAGCTCATAGTGATCATCACCACACTCTGCA----- 185
QY 116 GluProLeuGlnLeuHisValAspLysAlaValSerGlyLeuArgSerLeuThrLeu 135
|||||
DB 186 -----CTCCACCTCCACATAGAAAGCAGACATCGAATATTGCGAGCTCAAGCAGGTG 239
QY 136 LeuArgAlaLeuGlyAlaGlnLysGlnAlaLeuSerProAspAlaAlaSerAlaAla 155
|||||
DB 240 CTGCGACCTCCAGTATPAG-----GACTATATG 269
QY 156 ProLeuArgThrIleThrAlaAspThrPheArg-----LysLeuPheArg 170
|||||
DB 270 CCTACAAGTGTGGGACAGAGACACATGGAAGTCTTTCGGCTTCAAGACTCTTCAG 329
QY 171 ValITrSerAsnPhelLeuArgGlyLysLeuLysLeuThrArgGlyAlaGly 187

Db	330	GTCACATCAGCTTCTCGTGCGGAAAGACGCCCTTTTCTCCCAATGCT	380
RESULT 3	BE663803	444 bp mRNA linear	EST 25-APR-2001
LOCUS	BE663803		
DEFINITION	147806 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.		
ACCESSION	BE663803		
VERSION	BE663803.1 GI:10022630		
KEYWORDS	EST.		
SOURCE	cow.		
ORGANISM	Bos taurus		
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
Bovidae; Bovinae; Bos.			
REFERENCE	I (bases 1 to 444)		
AUTHORS	Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Pierce,G., Holt,I., Karameyehva,S., Liang,F., Quackenbush,J. and Keeler,J.W.		
TITLE	Sequence evaluation of four pooled-tissue normalized bovine CDNA libraries and construction of a gene index for cattle genome Res. 11 (4), 626-630 (2001)		
JOURNAL MEDLINE	21180013		
COMMENT	Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smithemail.marc.usda.gov Single pass sequencing. Bases called and alt-trimmed with phred v0.980904.e. Vector identified by cross_match with the minscore 18 and -minmatch 12 options. PCR Primers FORWARD: AGGAACAGCTATGCACAT BACKWARD: GTTTCCTCCACTCACGACG Plate: 51 row: H column: 3 Seq primer: ATTTAGGTGACACTYTAG. Location/Qualifiers 1..444 /organism="Bos taurus" /db_xref="taxon:9913" /clone_lib="MARC 4BOV" /tissue_type="pooled" /lab_host="DH10B" /note="Vector: pCMV SPORT6; Site.1: NotI; Site.2: SalI; Library made from pooled tissue from day 20 and day 40 embryos."		
FEATURES	source		
BASE COUNT	110 a 85 c 142 g 106 t	1 others	
ORIGIN			
Alignment Scores:			
Pred. No.:	2,83e-08	Length:	444
Score:	159.00	Matches:	28
Percent Similarity:	78.95%	Conservative:	2
Best Local Similarity:	73.68%	Mismatches:	8
Query Match:	15.93%	Indels:	0
DB:	10	Gaps:	0
US-09-813-775C-2 (1-193) x BE663803 (1-444)			
Oy	55	glycysalaglunhisccserleuansngluasniethrvalproaspthrlysalasn	74
Dd	152	ggctgtgcaggaaggcttcacatgatgattcatctgtgccagaccacaaggtttaac	211
Oy	75	phetyralatrplysarqmetclualatagglinglnalaivalgluvalttrpcin	92
Dd	212	tcttatgaccttgamaagagatggagtgcattcacgggttccttcccttggaa	265
RESULT 4	CNS02F50	968 bp DNA linear	GSS 13-MAY-2000
LOCUS	CNS02F50		

[illegible]

Insert. Size selection was performed to exclude fragments <1.5kb. Library constructed and donated by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTCTGCTTAAAGTGGC and 3' end primer CGACCTGCACCTCGACGACA.

BASE COUNT 156 a 172 c 144 g 139 t
ORIGIN

Alignment Scores:
Pred. No.: 0.00315 Length: 611
Score: 117.00 Matches: 46
Percent Similarity: 34.67% Conservative: 23
Best Local Similarity: 23.12% Mismatches: 49
Query Match: 11.72% Indels: 81
DB: 13 Gaps: 6

US-09-813-775c-2 (1-193) x BM315886 (1-611)

```
OY 2 G1yValHisGluCysProAlaTTrp-----LeuTrpLeuLeuSer 15
Db 233 GGAGTGGACAGTGTGT-----TGGCCAGGCTTTCATTAAGAGCTTGTGCTTACTG--- 283
OY 16 LeuLeuSerLeuProLeuGlyLeuProValLeuGlyAlaProProAlaGlyLeu11eCysasp 35
Db 284 CTGATGTGTCTGAGTGGAGCCGCTGCTCCCTCCATTAACCCCATCTGTGAC 343
OY 36 SerArgValLeuGluArgTyrLeuLeuGlu-AlaLysGluAlaGluAsn11eThrThgl 55
Db 344 CTGGCGGCTCCGACACATTTTCATACAGATATCTAATCAGACT----- 386
OY 55 YCysAlaGluHisCysSerLeuAsnGluAsn11eThrValProAspThrLysValAsnph 75
Db 386 ----- 386
OY 75 eTyrAlaTrpLysArgMetGluValArgGlnGlnAlaValGluValTrrpGlnGlyLeuAl 95
Db 386 ----- 386
OY 95 aLeuLeuSerGluAlaValLeuArgGlyGlnAlaLeuLeuValAsnSerSerGlnProTr 115
Db 386 ----- 386
OY 115 pGluProLeuGlnLeuHisValAspLysAlaValSerGlyLeuArgSerLeuThrThLe 135
Db 387 -GAAGTGTCTGCTCAGATAGATGCCATTAAGAAACATCCAGCAGATCAGACAGT 445
OY 135 uLeuArgAlaLeuGlyAlaGlnLysGluAla11eSerProProAspAlaAlaSerAlaAl 155
Db 446 GCTGCGAAGTCTCAGCATACCGAATATGTA-----CCTCCAACCCAGTAGTGAGA 496
OY 155 aProLeuArgThrThrAlaAspThrPheArgLysLeuPheArgValTyrSerAsnph 175
Db 497 AGAAGAGAGACACAGAAATATCTCGATCTCAGAGCTGTTTTCAGGTCCTGTCACACTT 556
OY 175 eLeuArgGlyLysLeuLysLeuTyrThrGlyGluAla-----CysArgThrGly 191
Db 557 TCTTCGGGGAAAGCGGCTGTGCTGCGCAATGCACTGTGTGTGCACAGGTT 611
```

RESULT 7
BG475544

LOCUS BG475544 967 bp mRNA linear EST 21-MAR-2001
DEFINITION 602491530F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4620438 5',
mRNA sequence.

ACCESSION BG475544
VERSION BG475544.1 GI:13407823

KEYWORDS
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 967)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC/DCPD/DRP

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
plate: LNCM381 row: 1 column: 07
High quality sequence stop: 525.
Location/Qualifiers
1. 967

FEATURES
source

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4620438"
/clone_id="NIH_MGC_20"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB1; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GCCACGAC(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 237 a 255 c 267 g 208 t  
ORIGIN
```

Alignment Scores:

Pred. No.: 0.639 Length: 967
Score: 99.50 Matches: 55
Percent Similarity: 37.17% Conservative: 29
Best Local Similarity: 24.34% Mismatches: 68
Query Match: 9.97% Indels: 74
DB: 12 Gaps: 12

US-09-813-775c-2 (1-193) x BG475544 (1-967)

```
OY 2 G1yValHisGluCysProAlaTTrpLeu---TrpLeuLeuSerLeuLeuSerLeuPro 20
Db 363 GGCTGTGCTGCTGCTGCTGCTTTCATCAGTGTG-----CTGCCA 401
OY 21 ---LeuGlyLeuProValLeu-----GlyAlaProProArg----- 31
Db 402 AGAGTTCACCTCCAGATCTCCAGGATCAGAGAGCCGCCCTCAGACAGATATTATC 461
OY 32 -----Leu11eCysAspSerArgValLeuGluArgTyrLeuLeu-----Glu 45
Db 462 CTTAACACTGTGCAACTCTGMAAGCTTGTCTGTCGCAAGAAATTTCTGATGATAG 521
OY 46 AlaLysGluAlaGluAsn11eThrThrGlyCysAlaGlnHisCysSerLeuAsnGluAsn 65
Db 522 GTCTTATAGTCCAAAGTGCCTAGCTGGAGATGTGCGAATGTG----- 566
OY 66 11eThrValProAspThrLysValAsnPheTyrAlaTrpLysArgMetGluValArgGln 85
Db 567 -----CCAGATGGGAGATCTCATCTGATCTGACAGAACAGATGCTCAGCAAAATAT 617
OY 86 GlnAlaValGluVal-----TrpGlnGlyLeuAlaLeuLeuSerGluAlaValLeu 102
Db 618 TCCCATGTCATTACCGTCGTCATTTGGGAAAGACTTGAGCGGAGAACCGGCAAGACTT 677
OY 103 ArgGlyGlnAlaLeuLeuValAsnSerSerGlnProTrpGluGlnLeuHisVal 122
Db 678 CGAGATCAACACACTTCAGAGATGGAGTGTGCATTC-----CACTGG 722
OY 123 AspLysAlaValSerGlyLeuArgSerLeuThrThrLeu-LeuArgAlaLeuGlyAlaG 142
Db 723 GAGCAGCAGCGTACAGACATTAATTCCTCCAAACCAATTAATTCAGACCGACAAAGA----- 777
```

OY 142 nlysgluAlaIleSerProProAspAlaIleSerAlaAlaProLeuArgThrIleThrAl 162
 DB 778 -----GGGGCAACCATCAGCCCGACATCATTCATTCACCT 812
 OY 162 aasprhrphearlysleuphearlyValTyrSerAspPheLeuArgGlyIlylsleuIysle 182
 DB 813 GGAAGAGTTTCACGAC-----TCCTTATGCAGGAGAACAGCATCAAT 854
 OY 182 u-----TCCTTATGCAGGAGAACAGCATCAAT 854
 DB 855 TACTAACATAGGCTGTGAACAACACTGTTCGACAACAGAGTCTGTTCGATTGAGTACCT 914
 OY 184 rcllygluaIaCysArg 189
 DB 915 TGCGAAGGCTGTGTAG 930
 RESULT 8
 BFL84120
 LOCUS 867 bp mRNA linear EST 31-OCT-2000
 DEFINITION 601843157F1 NIH_MGC_54 Homo sapiens cDNA clone IMAGE:4063767 5',
 mRNA sequence.
 ACCESSION BFL84120
 VERSION BFL84120.1 GI:11062479
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 867)
 NIH-MGC http://mhc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Clontech Laboratories, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLMC897 row: a column: 16
 High quality sequence stop: 659.
 FEATURES
 source
 Location/Qualifiers
 1..867
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4063767"
 /clone_lib="NIH_MGC_54"
 /tissue_type="from chronic myelogenous leukemia"
 /lab_host="DH10B (TI phage-resistant)"
 /note="Organ: bone marrow; Vector: pDNR-LIB (Clontech);
 Site_1: SfiI (ggcgccctcgccg); Site_2: SfiI (ggcgatctggcc
); Double-stranded cDNA was prepared from cell line RNA.
 5' and 3' adaptors were used in cloning as follows: 5'-
 adaptor sequence: 5'-CACGGCCATTATGACC-3' and 3' adaptor
 sequence: 5'-ATTCTAGAGCCGAGCGCGCATG-dt(30)BN-3'
 (where B = A, C, or G and N = A, C, G, or T). Average
 insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
 contained inserts by PCR. This library was enriched for
 full-length clones and was constructed by Clontech
 Laboratories (Palo Alto, CA)."
 BASE COUNT 223 a 193 c 233 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.06 Length: 867
 Score: 94.50 Matches: 26
 Percent Similarity: 50.00% Conservative: 17
 Best Local Similarity: 30.23% Mismatches: 34
 Query Match: 9.47% Indels: 9
 DB: 12 Gaps: 3

US-09-813-775C-2 (1-193) x BFL84120 (1-867)
 OY 7 ProAlaTrpLeuTrp---LeuLeuSerLeuSerLeuProLeuGlyLeuProVal 25
 DB 322 CCATTCCTTGTCTGGGAGAGTCGCCCTTCGCTTGGAATCATCTCTCTTTTCCCTGCA 381
 OY 26 LeuGlyAlaProProArgLeuIleCysAspSerArgValLeuGluArgTyrLeuLeuGlu 45
 DB 382 -----CTGGTTTGTGCCAGCGCTGCTCGCTGCTGCATTCATTCACG 423
 OY 46 AlAlaYgluaIaGluAsnIleThrTrpGlyCysAlaGluHisCysSerLeuAsnGluAsn 65
 DB 424 CTAGCGATGAGAACTCAGAAAGTCAGGCTTCATTCATTAATGATGATTCGCGGA 483
 OY 66 IleThrValProAspThrIlyrsValAsnPhetyrAlaTrpIlylsArgMetGluValArgGln 85
 DB 484 CTACAGGCTGCTGATACACAAATGATTTGTTTGTGG-----ATTGAAACCCAGAT 537
 OY 86 GlAlaValaIgluValTrp 91
 DB 538 AAGGAAAGATAGTTTGG 555
 RESULT 9
 BFL48123
 LOCUS 803 bp mRNA linear EST 12-SEP-2001
 DEFINITION 603278441F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5318674 5',
 mRNA sequence.
 ACCESSION BFL48123
 VERSION BFL48123.1 GI:15562359
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 803)
 NIH-MGC http://mhc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-r@mail.nih.gov
 Tissue Procurement: Lotmar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM11808 row: a column: 11
 High quality sequence stop: 802.
 FEATURES
 source
 Location/Qualifiers
 1..803
 /organism="Mus musculus"
 /strain="129, C57BL/6J FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5318674"
 /clone_lib="NCI_CGAP_Mam3"
 /tissue_type="tumor, gross tissue"
 /dev_stage="10 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigators
 providing samples: Lotmar Hennighausen/Chu-Xia Deng, NIH
 Reference for transgenic model: Xu et al., Nature Genetics
 22, 37-43 (1999)."
 BASE COUNT 171 a 252 c 199 g 181 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.1 Length: 803
 Score: 94.00 Matches: 51
 Percent Similarity: 37.44% Conservative: 25

OY 93 -----GlyLeuAlaLeuLeuSerGluAlaVal----- 101
 Db 417 AAGGGGAGAGTACGCCCTGAGGCGACATGAGGCGACGCCGCCACGTAGA 358
 OY 102 -----LeuArgGlyGlnAlaLeuLeuValAsnSerSer 112
 Db 357 GGAACCTGTTTCGGGCGAGCTGGGTGTCTCCCGACAGCAGCAGCTTTCGAATCCAAAGC 298
 OY 113 GlnProtrpGluPro--LeuGlnLeuHisValAspLysAlaValSerGlyLeu-----A 130
 Db 297 CAAGCTCAGAAACACAGCAGCAGAGTCCGGGCGCCCTAGGCGCTCCCGCTTAAATTCGCC 238
 OY 130 rgSerLeuThrThrLeuLeuArgAlaValGlnGlyAlaGlnLysGluAlaLysSerProPro 149
 Db 237 GAGGTCCAGACTCAAGATAGCACACATTAGGCGCTCGAAGCAGAGCAATCCAGCCTCCC 179

RESULT 13

BJS28644/c

LOCUS BJS28644 MF01SSB 639 bp mRNA linear EST 09-AUG-2002
 DEFINITION mRNA sequence.
 ACCESSION BJS28644
 VERSION BJS28644.1 GI:22187456
 KEYWORDS EST.
 SOURCE Japanese medaka.
 ORGANISM Oryzias latipes.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
 Belontiiformes; Adriantichthyidae; Oryziatidae; Oryzias.
 1 (bases 1 to 639)
 AUTHORS Kohara,Y., Shin,I.T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
 TITLE Medaka EST Project in Takeda's lab
 JOURNAL Unpublished (2001)
 COMMENT Contact: Tadasu Shin-1
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6855
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp.

FEATURES

source

1..639
 location/Qualifiers
 /organism="Oryzias latipes"
 /strain="Hd-r"
 /db_xref="taxon:8090"
 /clone="MF01SSB033E16"
 /clone_lib="MF01SSB cDNA"
 /sex="mixture of female and male"
 /tissue_type="whole embryo"
 /dev_stage="segmentation stage 20 - 25"
 BASE COUNT 175 a 137 c 228 g 97 t 2 others
 ORIGIN

Alignment Scores:
 Pred. No.: 3.35 Length: 639
 Score: 91.00 Matches: 51
 Percent Similarity: 37.31% Conservative: 21
 Best Local Similarity: 26.42% Mismatches: 60
 Query Match: 9.12% Indels: 61
 DB: 13 Gaps: 9

US-09-813-775C-2 (1-193) x BJS28644 (1-639)

OY 5 GluGys-----ProAlaTrp-----LeuThrLeuLeuLeuSerLeuLeu 17
 Db 546 GAATGTTGAACCTCCACCTTGGCCGCTTCCTTGCCTTCATGCGTTCATTCGTCACAG 487
 OY 18 SerLeuProLeuGlyLeuProValLeuGlyAlaProProArgLeuLeuLecysAspSerArg 37
 Db 486 TCATTCCTTGGGCGCTTCCTTCCTTCACCTTCCTTCACCTTCAT----- 445

OY 38 ValLeuGluArgTyrLeuLeuGluAlaLysGluAlaLysAsnIleThrThrGlyCysAla 57
 Db 445 ----- 445
 OY 58 GluHisCysSerLeuAsnGluAsnIleThrValProAspThrLysValAsnPheTyrAla 77
 Db 444 TTCCTCTTGTCTCAGAGTACAGCAGCTGATGTTCT----- 409
 OY 78 TrpLysArgMetGluValArgGlnGlnAlaValGluValTrp-GlnGlyLeuAlaLeuLe 97
 Db 408 -----CTCCGTCAGAGATTCCTTCAG---TGGTGTGATTTGGTCAAGCT 367
 OY 97 uSerGluAlaValLeuArgGlyGlnAlaLeuLeuValAsnSerSerGlnProtrpGluPro 117
 Db 366 CATTCAGTTCATCCTTCACGCT---GCCCAAGTTGTAGATTCACATTCCTCCCAAGCTTCT 310
 OY 117 oleuGlnLeuHis-----ValAspLysAlaValSerGlyLeuArgSerLeuThrLe 135
 Db 309 CCTACCTTTCATGATGATTTGTACACTGCGGTGTGTGTCAAAATTCACGCTTTCGCCGGG 250
 OY 135 uLeuArgAlaLeuGlyAlaGlnLysGluAlaLysSerProPro-----AspAlaAl 152
 Db 249 AGTCAAGCCGTCACCTGCGACCCATACCGGACCCCTGCCCCGGCGCTGC 190
 OY 152 aSerAlaAlaProLeuArgThrIleThr-----AlaAspThrph 165
 Db 189 CACCGCTCCCTCATGAGCGCGCATCCGATGCGGTTCCTCCACTGAGATTCACACCTT 130
 OY 165 eaArgLysLeuPheArgValLysSerAsnPheLeuArg 177
 Db 129 CAGTCTTCTTACAGCATGCCACGACGAGCTGCTCAAA 93

RESULT 14

BE618292/c

LOCUS BE618292 912 bp mRNA linear EST 20-OCT-2000
 DEFINITION 601462670F1 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:3866248 5',
 mRNA sequence.
 ACCESSION BE618292
 VERSION BE618292.1 GI:9889230
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 912)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 plate: LLAW610 row: k column: 17
 High quality sequence stop: 671.
 location/Qualifiers
 1..912
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3866248"
 /clone_lib="NIH_MGC-67"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pCMV-Sport6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo
 Average insert size 1.75 kb. Library constructed by Life
 Technologies."

FEATURES

source

BASE COUNT 180 a 316 c 264 g 152 t
 ORIGIN

Db 409 TCCTGACCTTATCCCTCCTCGGACAGCTTTCGTGGAGAGTTGACCTCCTCTTTGGGGGCC 548

Oy 166 ILeserProProspAlaAlaSerAlaAlaProLeuArgThrIleThr 161
::: ||| ||| ||| ||| |||

Db 549 CTGCACCGTCTTAAGGAACCCAGGGCCAGGACACAGCTTCACAGGAGACC 596

Search completed: January 8, 2003, 07:50:49
Job time : 1644.26 secs

	Type	L #	Hits	Search Text	DBs	Time Stamp	Comments	Error Definition	Errors
1	BRS	L1	5796	erythropoietin	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/01/08 19:34			0
2	BRS	L2	3	1 same chimpanzee	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/01/08 19:36			0
3	BRS	L3	0	2 same chimeric	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/01/08 19:37			0
4	BRS	L4	1	desauvage adj frederic.in.	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/01/08 19:38			0
5	BRS	L5	14	henner adj dennis.in.	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/01/08 19:39			0
6	BRS	L6	4	5 and 1	USPAT; US-PGPUB; EPO; JPO; DERWENT	2003/01/08 19:39			0

=> file medline caplus biosis embase scisearch agricola
COST IN U.S. DOLLARS
SINCE FILE
ENTRY
SESSION
TOTAL 0.21

FILE 'MEDLINE' ENTERED AT 19:45:10 ON 08 JAN 2003
FILE 'CAPLUS' ENTERED AT 19:45:10 ON 08 JAN 2003
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2003 AMERICAN CHEMICAL SOCIETY (ACS)
FILE 'BIOSIS' ENTERED AT 19:45:10 ON 08 JAN 2003
COPYRIGHT (C) 2003 BIOLOGICAL ABSTRACTS INC. (R)
FILE 'EMBASE' ENTERED AT 19:45:10 ON 08 JAN 2003
COPYRIGHT (C) 2003 Elsevier Science B.V. All rights reserved.
FILE 'SCISEARCH' ENTERED AT 19:45:10 ON 08 JAN 2003
COPYRIGHT (C) 2003 Institute for Scientific Information (ISI) (R)
FILE 'AGRICOLA' ENTERED AT 19:45:10 ON 08 JAN 2003

=> s erythropoietin
69950 ERYTHROPOIETIN

=> s 11 (p) chimpanzee
5 L1 (P) CHIMPANZEE

=> duplicate remove
ENTER L# LIST OR (END):12
DUPLICATE PREFERENCE IS 'MEDLINE', CAPLUS'
KEEP DUPLICATES FROM MORE THAN ONE FILE? Y/(N):n
PROCESSING COMPLETED FOR L2
L3 5 DUPLICATE REMOVE L2 (0 DUPLICATES REMOVED)
=> d 13 1-5 1b1b abs
L3 ANSWER 1 OF 5 CAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2002:736286 CAPLUS
DOCUMENT NUMBER: 137:257948
TITLE: ***Chimpanzee***
- immunoadhesins for use in regulating erythropoiesis
INVENTOR(S): Desauvage, Frederic; Henner, Dennis J.
PATENT ASSIGNEE(S): Genentech, Inc., USA
SOURCE: PCT Int. Appl., 120 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO. KIND DATE APPLICATION NO. DATE

WO 2002074807 A2 20020926 WO 2002-US4773 20020214

W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN,

CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, FI, GB, GD, GE, GH, GM, GR, HU, ID, IL, IN, IS, JP, KE, KG, KP,

KR, KZ, LC, LK, LR, LS, LT, LV, MA, MD, MG, MK, MN, MW, MX, MY, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW, AM, AZ,

BY, BG, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG,

PRIORITY APPLN. INFO.:
US 2001-813775 A 20010320
The present invention is directed to immunoadhesins comprising
chimpanzee
erythropoietin (CHEPO) polypeptides. The

immunoadhesins have an enhanced in vivo half-life compared to the corresponding CHEPO polypeptide and retains CHEPO bio. activity. Also provided herein are nucleic acid mols. encoding such immunoadhesins, vectors and host cells comprising those nucleic acid sequences, and methods using and compns. comprising the CHEPO immunoadhesins.

L3 ANSWER 2 OF 5 CAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2000:814616 CAPLUS
DOCUMENT NUMBER: 134:1044
TITLE: ***Chimpanzee***
INVENTOR(S): Desauvage, Frederic; Henner, Dennis J.
PATENT ASSIGNEE(S): Genentech, Inc., USA
SOURCE: PCT Int. Appl., 109 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2000068376	A1	20001116	WO 2000-US12370	20000505

W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, GU, HR, HU, ID, IL, IN, IS, JP, KE, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM
RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, BF, BJ, CF, CG, CI, CM, GN, GW, ML, MR, NE, SN, TD, TG
EP 117285 A1 20020206 EP 2000-928879 20000505
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO
JP 2002543784 T2 20021224
JP 2000-616342 20000505
US 1999-307307 A 19990507
US 2000-287594P P 20000328
US 2000-552265 A 20000419
WO 2000-US12370 W 20000505

PRIORITY APPLN. INFO.:
AB The present invention is directed to novel ***chimpanzee***
erythropoietins and to nucleic acid mols. encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric proteins comprising the ***erythropoietins*** of the present invention fused to heterologous protein sequences, antibodies which bind to the proteins of the invention, and to methods for producing the proteins of the invention. Agonists and antagonists of ***erythropoietin*** may be used to modulate erythropoiesis.
REFERENCE COUNT: 8
THERE ARE 8 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L3 ANSWER 3 OF 5 CAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 1999:708637 CAPLUS
DOCUMENT NUMBER: 131:335812
TITLE: Monoclonal antibodies with reduced immunogenicity
INVENTOR(S): Taylor, Alexander H.
PATENT ASSIGNEE(S): Smithkline Beecham Corporation, USA
SOURCE: PCT Int. Appl., 122 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9955369	A1	19991104	WO 1999-US9131	19990428

W: CA, JP, US
RM: AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE

CA 2327505 AA 19991104 CA 1999-2327505 19990428
EP 1073464 AI 200104 EP 1999-918862 19990428
R: BE, CH, DE, ES, FR, GB, IT, LI, NL
JP 2002512776 T2 20020508 JP 2000-545566 19990428
US 2002062009 AI 20020523 US 1998-83367P P 19980428
US 1999-300970 B1 19990428 US 1999-US9131 W 19990428
Antibodies having reduced immunogenicity comprising non-human
antigen-specific donor antibody and non-human primate acceptor framework.
Random cDNA cloning and sequence anal. of VH and V.kappa. of
chimpanzee and cynomolgus were performed, and engineered murine
monoclonal antibodies (contg. ***chimpanzee*** or cynomolgus
framework) specific for interleukin 5, integrin, and
erythropoietin receptor were prepd. for treating asthma, vascular
diseases and hematopoietic disorders resp.
REFERENCE COUNT: 1 THERE ARE 1 CITED REFERENCES AVAILABLE FOR THIS
RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L3 ANSWER 4 OF 5 CAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 1970:505985 CAPLUS

DOCUMENT NUMBER: 73:105985

TITLE:

Anti-human follicle stimulating hormone; complete
cross-reactivity with three other human glycoprotein
tropic hormones, luteinizing hormone, human chorionic
gonadotropin, and thyrotropin
AUTHOR(S): Rosen, Saul W.; Schlaft, Sheldon; Roth, Jesse
Nat. Inst. Arthritis and Metab. Dis., Nat. Inst. of
Health, Bethesda, MD, USA
Protein Polypeptide Horm., Proc. Int. Symp. (1969),
Meeting Date 1968, 396-7. Editor(s): Margoulies, M.
Excerpta Med. Found.: Amsterdam, Neth.

DOCUMENT TYPE: Conference

LANGUAGE: English

AB

Addn. of excess rabbit or guinea pig anti-human FSH sera to FSH-131I pptd.
70-85% of the radioactivity; nonimmune sera pptd. 6% of the radioactivity.
Excess unlabeled FSH displaced 95% of the antibody-bound radioactivity
into the supernatant. Displacement curves of FSH-131I from anti-FSH by
various FSH preps. were superimposable when the abscissas were adjusted
for varying FSH content. Plasma from postmenopausal patients, patients
with gonadal insufficiency due to the Turner syndrome or Klinefelter
syndrome, and patients with gonadotropin-producing bronchogenic carcinomas
showed curves of displacement superimposable over a 4-20 fold range of
dilin. Human LH, HCG (human chorionic gonadotropin), and TSH displaced
FSH-131I from anti-FSH. LH-131I and HCG-131I were bound as strongly as
FSH-131I by anti-FSH sera and the curves of displacement of LH-131I from
anti-FSH by LH and SFH were quant. identical. Human growth hormone,
bovine ACTH, ovine FSH, human ***erythropoietin***, arginine
vasotocin, human .alpha.-acid glycoprotein, bovine thyroglobulin, human
ceruloplasmin, bovine fetuin, human orosomucoid, and blood group A and B
substances did not displace FSH-131I from anti-FSH. Plasma from rabbits,
guinea pigs, dogs, cats, hogs, goats, sheep, burros, and horses, including
some aged and some castrate females, failed to displace FSH-131I.
Displacement of human FSH-131I from antihuman FSH was produced by sera
from the higher apes (***chimpanzee***, gorilla, orangutan, gibbon)
but not by sera from rhesus, pigtail, Java, or African green monkeys, or
from black apes, baboons, or drills. These data suggested that the
pituitary glycoprotein tropic hormones and chorionic gonadotropin in man,
and possibly in the social apes, share common antigenic sites.

L3 ANSWER 5 OF 5 MEDLINE

ACCESSION NUMBER: 68092392 MEDLINE

DOCUMENT NUMBER: 68092392 Pubmed ID: 6073630

TITLE:

Erythropoietin response to renal stress in the
chimpanzee.

AUTHOR:

Murphy G P, Mirand E A, Wade J C, Melby E C
INVESTIGATIVE UROLOGY, (1967 Nov) 5 (3) 234-8.
Journal code: 0374747. ISSN: 0021-0005.

PUB. COUNTRY: United States

DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)

LANGUAGE: English

FILE SEGMENT:
ENTRY MONTH:
ENTRY DATE:

Priority Journals
196802
Entered STN: 19900101
Last updated on STN: 19900101
Entered Medline: 19680215

=> d h i s

(FILE 'HOME' ENTERED AT 19:44:43 ON 08 JAN 2003)

FILE 'MEDLINE', CAPLUS, BIOSIS, EMBASE, SCISEARCH, AGRICOLA' ENTERED AT
19:45:10 ON 08 JAN 2003

L1 69950 S ERYTHROPOIETIN
L2 5 S L1 (P) CHIMPANZEE
L3 5 DUPLICATE REMOVE L2 (0 DUPLICATES REMOVED)

=> s chimeric
L4 105699 CHIMERIC

=> s l3 (p) l4

PROXIMITY OPERATOR LEVEL NOT CONSISTENT WITH
FIELD CODE - 'AND' OPERATOR ASSUMED 'L27' (P) L18'

PROXIMITY OPERATOR LEVEL NOT CONSISTENT WITH
FIELD CODE - 'AND' OPERATOR ASSUMED 'L29' (P) L19'

PROXIMITY OPERATOR LEVEL NOT CONSISTENT WITH
FIELD CODE - 'AND' OPERATOR ASSUMED 'L31' (P) L20'

PROXIMITY OPERATOR LEVEL NOT CONSISTENT WITH
FIELD CODE - 'AND' OPERATOR ASSUMED 'L33' (P) L21'

L5 2 L3 (P) L4

=> d l5 1-2 l5b abs

L5 ANSWER 1 OF 2 CAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2002:736286 CAPLUS
DOCUMENT NUMBER: 137:257948

TITLE: Chimpanzee erythropoietin (CHEPO) - immunoadhesins for

INVENTOR(S): Desauvage, Frederic; Henner, Dennis J.
PATENT ASSIGNEE(S): Genentech, Inc., USA

SOURCE: PCT Int. Appl., 120 pp.

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO. KIND DATE APPLICATION NO. DATE

WO 2002074807 A2 20020926 WO 2002-US4773 20020214

W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, FL, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KP, KR, KZ, LC, LK, LR, LS, LT, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW, AM, AZ, BY, BG, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG

PRIORITY APPLN. INFO.: US 2001-813775 A 20010320

AB The present invention is directed to immunoadhesins comprising chimpanzee erythropoietin (CHEPO) polypeptides. The immunoadhesins have an enhanced in vivo half-life compared to the corresponding CHEPO polypeptide and retains CHEPO biol. activity. Also provided herein are nucleic acid moles encoding such immunoadhesins, vectors and host cells comprising those nucleic acid sequences, and methods using and comps. comprising the CHEPO immunoadhesins.

L5 ANSWER 2 OF 2 CAPLUS COPYRIGHT 2003 ACS
ACCESSION NUMBER: 2000:814616 CAPLUS
DOCUMENT NUMBER: 134:1044

TITLE: Chimpanzee erythropoietin and cDNA and methods of

modulate erythropoiesis

Desauvage, Frederic; Henner, Dennis J.

Genentech, Inc., USA

PCT Int. Appl., 109 pp.

CODEN: PIXXD2

Patent

English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.

KIND DATE

APPLICATION NO. DATE

WO 2000068376 A1 20001116 WO 2000-US12370 20000505

AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, GU, HA, HE, HI, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, ST, TM, TR, TT, TZ, UA, UG, US, VZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM

RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GN, GW, ML, MR, NE, SN, TD, TG

EP 117285 A1 20020206 EP 2000-928879 20000505

JP 2002543784 T2 20021224 JP 2000-616342 20000505

US 1999-307307 A 19990507 US 2000-287594P P 20000328 US 2000-552265 A 20000419 WO 2000-US12370 W 20000505

AB The present invention is directed to novel ***chimpanzee***

erythropoietins and to nucleic acid mols. encoding those

polypeptides. Also provided herein are vectors and host cells comprising

those nucleic acid sequences, ***chimeric*** proteins comprising the

erythropoietins of the present invention fused to heterologous

protein sequences, antibodies which bind to the proteins of the invention,

and to methods for producing the proteins of the invention. Agonists and

antagonists of ***erythropoietin*** may be used to modulate

erythropoiesis.

REFERENCE COUNT: 8

THERE ARE 8 CITED REFERENCES AVAILABLE FOR THIS

RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

=> d his

(FILE 'HOME' ENTERED AT 19:44:43 ON 08 JAN 2003)

FILE 'MEDLINE', CAPLUS, BIOSIS, EMBASE, SCISEARCH, AGRICOLA, ENTERED AT

19:45:10 ON 08 JAN 2003

L1 69950 S ERYTHROPOIETIN

L2 5 S L1 (P) CHIMPANZEE

L3 5 DUPLICATE REMOVE L2 (0 DUPLICATES REMOVED)

L4 105699 S CHIMERIC

L5 2 S L3 (P) L4

L6 0 EPILOPE ADJ TAG

=> s epitope tag

L7 1837 EPILOPE TAG

=> s 17 and 15

L8 0 L7 AND L5

=> s immunoglobulin

L9 2 IMMUNOGLOBULIN

IN IS NOT A RECOGNIZED COMMAND

The previous command entered was not recognized by the system.

For a list of commands available to you in the current file, enter
"HELP COMMANDS" at an arrow prompt (=>).

=> s immunoglobulin
L10 719573 IMMUNOGLOBULIN

=> s 15 (p) 110
PROXIMITY OPERATOR LEVEL NOT CONSISTENT WITH
FIELD CODE - 'AND' OPERATOR ASSUMED 'L28 (p) L66'
PROXIMITY OPERATOR LEVEL NOT CONSISTENT WITH
FIELD CODE - 'AND' OPERATOR ASSUMED 'L30 (p) L67'
PROXIMITY OPERATOR LEVEL NOT CONSISTENT WITH
FIELD CODE - 'AND' OPERATOR ASSUMED 'L32 (p) L68'
PROXIMITY OPERATOR LEVEL NOT CONSISTENT WITH
FIELD CODE - 'AND' OPERATOR ASSUMED 'L34 (p) L69'

=> s 19g or 19g1
L12 340314 IGG OR IGG1

=> s 112 and 15
L13 0 L12 AND L5

=> d his

(FILE 'HOME' ENTERED AT 19:44:43 ON 08 JAN 2003)

FILE 'MEDLINE', CAPLUS, BIOSIS, EMBASE, SCISEARCH, AGRICOLA' ENTERED AT
19:45:10 ON 08 JAN 2003

L1 69950 S ERYTHROPOIETIN
L2 5 S L1 (p) CHIMPANZEE
L3 5 DUPLICATE REMOVE L2 (0 DUPLICATES REMOVED)
L4 105699 S CHIMERIC
L5 2 S L3 (p) L4
L6 0 S EPILOPE ADJ TAG
L7 1837 S EPILOPE TAG
L8 0 S L7 AND L5
L9 2 S IMMUNOGLOBU
L10 719573 S IMMUNOGLOBULIN
L11 0 S L5 (p) L10
L12 340314 S IGG OR IGG1
L13 0 S L12 AND L5

=> log y
COST IN U.S. DOLLARS

SINCE FILE ENTRY TOTAL
41.01 41.22
FULL ESTIMATED COST

DISCOUNT AMOUNTS (FOR QUALIFYING ACCOUNTS)
SINCE FILE ENTRY TOTAL
-3.91 -3.91
CA SUBSCRIBER PRICE

STN INTERNATIONAL LOGOFF AT 19:50:27 ON 08 JAN 2003

